UNIVERSITY of the WESTERN CAPE

Faculty of Science
Department of Statistics

Statistics Distribution Theory
STA 211

First Semester 2020

Contents

Part A. General Information 2

1. Teaching Staff and Contact Times 2
2. Module Overview 2

Part B. Teaching and Learning 4

3. Teaching and Learning objectives 4
4. Module Schedule 5
5. Materials 6
6. Graduate Attributes, Learning Outcomes and Assessment 7
7. Evaluation of the Teaching and Learning 8

Part C. General Information 8

8. Academic Honesty and Discipline 8

Welcome to the Department of Statistics and STA 211. This is a first semester 20 credit course. Students interested in studying Statistics in their third year will also need to register for STA 221 in the second semester.

Part A. General Information

1. Teaching Staff information

Module Coordinator: Mrs. Rechelle Jacobs

Lecturer:	
Name	**Mrs. R Jacobs**
Room and Building	Room 1.91, CAMS Building
Phone Number	021 959 9565
Email	rejacobs@uwc.ac.za
Consultation Times	Strictly by appointment due to access control

Class Times

- Mondays, 08h30-09h15 in N22
- Tuesdays, 11h15-12h00 in N22
- Wednesdays, 08h30-09h15 in N22
- Wednesdays, 12h10-12h55 in N22

Tutorial Times

- Thursdays, 11h15-12h55 (period 4 and 5) in Labs K1 and K2 (CAMS Building)

Practical Demonstrator:	
Name	**Mr. M Valentine**
Room and Building	Tech Room K2, CAMS Building
Phone Number	021 959 1613
Email	mvalentine@uwc.ac.za
Consultation Times	Tuesdays, 9h30-11h20 (Lab K2 Tech Room K2)
Practical Times	Thursdays, 14h00-16h35 (period 6, 7 and 8) in Labs K1 and K2 (CAMS Building)

2. Module Description

Distribution theory as taught in STA 211 is concerned with Statistics, in both theory and application. The key aims are to convey a thorough understanding of the fundamentals of random variables and their distributions and their role in inference and more broadly, research. The course builds on a brief introduction to inference and elementary Statistics. It is assumed that students are familiar with these ideas and so the course is fast paced but fair.

Faculty of Science Module Descriptor:

Faculty	Natural Science	
Home Department	Statistics	
Module Topic	Distribution Theory	
Generic module name	Statistics 211	
Alpha-numeric code	STA211	
NQF Credit Value	20	
Duration	Semester	
Proposed semester to be offered. (For Calendar Groups)	First semester	
Programmes in which the module is offered.	Mathematical and Statistical Sciences, Computer Science, B.Sc (General), B.Com (General)	
NQF level	5	
Year Level	2	
Main Outcomes	To be able to gain insight into and apply - Probability theory, - Discrete and continuous probability distributions, - Moments and moment generating functions, - Multivariate Probability distributions, - Develop statistical computer literacy skills.	
Main Content	*Distribution theory:* - Definition of statistical terms; - Probability theory; - Discrete and continuous probability distributions; - Moments and moment generating functions; - Multivariate Probability distributions; - Manipulating and summarizing data with reports and graphs.	
Pre-requisites	MAT105/(103+Registered for 104)/MAM(151+152)/115/150/126/127/ and STA111/125/141/142/151/BUS131/132 or equivalent	
Co-requisites	None	
Prohibited Combinations	None	
A. Breakdown of Learning Time	Current Hours	B. Time-table Requirement per week • Lectures p.w. 4 x 45 minutes • Tutorials p.w. 2 x 45 minutes • Practicals p.w. 3 x 45 minutes
Contact with lecturer / tutor:	42	
Assignments & tasks:	48	
Assessment:	14	
Practical's:	31	
Self- study:	65	
Other: Please specify	0	
Total Learning Time:	200	
Time-table Requirement per week	Lectures p.w. 4 x 45 minutes	
	Practicals p.w. 3 x 45 minutes (CAMS)	
	Tutorials p.w. 2 x 45 minutes (CAMS)	
Methods of Student Assessment	Tests, assignments, tutorials and practical's: 50% Final examination: 50%	
Assessment Module type	CFA	

At the end of the course you should be able

- Understand probability and probability distributions
- Calculate probabilities associated with simple, realistic and complex experiments,
- Use the methods for finding probability distributions for functions of random variables and to
- Understand how these functions aid in evaluating the goodness of statistical procedures
- Understand the role probability plays in making inferences and to put together all the theory and techniques you have learnt to solve practical problems ,
- Use MS Excel® Statistical functions and understand how to solve Statistical problems using MS Excel®,
- Examine data to be used with SAS®, read data into SAS®, code DATA and PROC steps in a SAS® program and interpret a SAS® log.

Part B. Teaching and Learning

3. Teaching and Learning Objectives

Students are expected to:

- Attend all lectures, tutorials and practical sessions;
- Come prepared to lectures;
- Complete weekly tutorials to be tested on every week;
- Complete and submit practical assignments on time.

Lectures:
Lectures are in N22. Students are required to **study the relevant material before attending** class.

Tutorials:
Weekly tutorials are in labs K1 and K2. Each week's tutorial questions are indicated under the "tutorial" column below. ***The tutorial exercises will not be handed in to be marked but rather a tutorial test on the due tutorial will be given. One theory question and one question similar to one of the tutorial exercises will be in the tutorial test. The tutorial test mark will then be the mark achieved for the tutorial.***

Practical's:
Practical's are computer based in the K1 and K2 labs. The schedule is below.

Tests:
The tests will consist of theory, insight questions and problems. Approximately Sixty percent (60%) of the test will be theory questions, questions similar to those in the tutorial exercises and the rest (approximately 40%) will consist of self-study and insight questions and applications or not seen before problems.

4. Module Schedule

Week	Theory	Assessment
Week 1 3/2/20	Chapter 1.1 to 1.6 (Self-Study) Chapter 2.1 to 2.8 (Revision sections: 2.1-2.3 & 2.6-2.8)	Revision exercises: All of Chapter 1 & 2.1; 2.2; 2.3; 2.6; 2.7; 2.8 (exclude applet exercises). **Tutorial Class date: 6/2** (Chapters 2.1 to 2.8) **Tutorial 1 exercises:** 2.12; 2.23; 2.32; 2.33; 2.76. Self-study exercises: 2.10; 2.19; 2.25; 2.28; 2.77.
Week 2 10/2/20	Chapter 2.9 to 2.13. Chapter 3.1 to 3.2	**Tutorial 2-Exercises:** 2.117; 2.125; 2.129; 2.136; 2.173. Self-study exercises: 2.110; 2.120; 2.130; 3.1; 3.5. **Tutorial 1 - Test date: 13/2**
Week 3 17/2/20	Chapter 3.3 to 3.6	**Tutorial 3 exercises:** 3.17; 3.64; 3.65; 3.86; 3.87. Self-study exercises: 3.14; 3.18; 3.58; 3.33; 3.101. **Tutorial 2 - Test date: 20/2**
Week 4 24/2/20	Chapter 3.7 to 3.11	**Tutorial 4 exercises:** 3.113; 3.135; 3.145; 3.146; 3.167 Self-study exercises: 3.121; 3.147; 3.148; 3.164; 3.171. **Tutorial 3 - Test date: 27/2**
26/2/20	**Test 1 - Chapter 2 & 3**	
Week 5 2/3/20	Chapter 4.1 to 4.4	**Tutorial 5 exercises:** 4.16; 4.20; 4.26; 4.41; 4.53 Self-study exercises: 4.1; 4.13; 4.21; 4.30; 4.51. **Tutorial 4 - Test date: 5/3** **Hand Written Theory Report 1: Chapters 2 & 3 - Due date: 5/3**
Week 6 9/3/20	Chapter 4.5 to 4.7	**Tutorial 6 exercises:** 4.99 a; 4.111; 4.112; 4.127; 4.130; 4.134a Self-study exercises: 4.60; 4.72; 4.80; 4.97; 4.109. **Tutorial 5 - Test date: 12/3**
Week 7 16/3/20	Chapter 4.8 to 4.10	**Tutorial 7 exercises:** 4.144; 4.146; 4.196; 4.199; 4.200. Self-study exercises: 4.136; 4.140; 4.141; 4.143; 4.194. **Tutorial 6 - Test date: 19/3**
Mid-Term Break: Wednesday 21/3 (Human Rights Day)		
Week 8 30/3/20	Chapter 4 - Revision	**Tutorial 7 - Test date: 2/4**
1/4/20	**Test 2 - Chapter 4**	
Week 9 6/4/20	Chapter 5.1 to 5.3	**Tutorial 8 exercises**: 5.9; 5.27; 5.33; 5.34; 5.36. Self-study exercises: 5.8; 5.15; 5.16; 5.26; 5.35. **Hand written Theory Report 2: Chapters 4&5 - Due date: 9/4**
Holiday: Friday 10/4 - Good Friday & Monday 13/4 - Family Day		
Week 10 14/4/20	Chapter 5.4 to 5.7	**Tutorial 9 exercises:** 5.60; 5.65; 5.74; 5.80; 5.100. Self-study exercises: 5.49; 5.78; 5.87, 5.92; 5.93. **Tutorial 8 - Test date: 16/4**
Week 11 20/4/20	Chapter 5.8 to 5.9	**Tutorial 10 exercises:** 5.108; 5.114; 5.121; 5.122; 5.124. Self-study exercises: 5.106; 5.119; 5.123; 5.125; 5.126. **Tutorial 9 - Test date: 23/4**
Week 12 27/4/20	Chapter 5.10 to 5.11	**Tutorial 11** exercises**:** 5.128; 5.219; 5.130; 5.131; 5.133; 5.136; 5.138; 5.140; 5.141; 5.142. **Tutorial 10 - Test date: 30/4**
Holiday: 27/4 - Workers Day & 1/5 - May Workers		
Week 13 4/5/20	Chapter 5 Revision	**Tutorial Exercise Revision**
6/5/20	**Test 3 - Chapter 5**	
11/5/20	**Sick Test - Chapters 3 to 5**	
Week 14 12-15/5	Theory Revision	Past Exam Papers

Practical's: Schedule

Dates	Practical Activity	Assessment
06-Feb-20	• Summarizing data by using descriptive statistics. • Using histograms and descriptive statistics	**Excel Assignment** Based on the work covered in the session. Due:13/02/2020
13-Feb-20	• Using pivot tables, pivot charts and slicers to describe data. • Estimating straight line relationship (Linear Regression)	**Excel Assignment** Based on the work covered in the session. Due:20/02/2020
20-Feb-20	• Modelling Exponential growth • The Power curve • ANOVA (Analysis of Variance)	**Excel Assignment** Based on the work covered in the session. Due: 27/03/2020
27-Feb-20	• Binomial Distribution • Poisson Distribution • Normal Distribution • T-Distribution • F- Distribution • Chi-Square Distribution	**Excel Assignment** Based on the work covered in the session. Due:05/03/2020
05-Mar-20	• **SAS - Chapter 1:** Getting Started **Chapter 2:** Getting Data into SAS	
12-Mar-20	• **SAS Practical Test 1 - Chapter 1 & 2 (Theory Test)**	
19-Mar-20	• **SAS – Chapter 3:** Reading, Writing and Importing Data	
Mid-Term Break 20 - 29 Mar-20		
02-Apr-20	• **Chapter 4: Preparing Data for Analysis**	
09-Apr-20	• **SAS Practical Test 2 - Chapters 3 & 4**	
16-Apr-20	• **SAS – Chapter 5:** Preparing to use SAS Procedures **Chapter 6:** Evaluating Quantitative Data	
23-Apr-20	• **SAS Practical Test 3 - Chapter 5 & 6**	
30-Apr-20	• **SAS – Chapter 7:** Analyzing Counts of Tables **Chapter 8:** Comparing Means using T-Tests	
7-May-20	• **SAS Practical Test 4 – Chapter 7 & 8**	

5. Materials

Prescribed Textbook: *Mathematical Statistics with Applications* by Wackerly, Mendenhall and Scheaffer 7th Edition

6. Graduate Attributes, Learning Outcomes and Assessment

UWC Graduate Attributes	**Learning outcomes**	**Teaching/Learning activities**	**Assessment tasks and criteria**		
Inquiry--focused	Apply statistics to simple and complex experiments.	Class discussion; In-class exercises; Pre-reading and preparation.	Semester test and final exam questions.	Weekly tutorial exercises.	Using EXCEL and SAS to assist in problem solving.

	Solve quantitative statistical problems.	Tutorial exercises; Computer analysis; Assignments;	Semester test and final exam questions.	Weekly tutorial exercises.	Statistical analysis reports.
Critically and relevantly literate	Conduct research using the library, the web and other sources of information.	Research on identified Topics; Oral presentation/discussion; Practical reports.	Semester test and final exam questions.		Statistical analysis reports.
	Reference sources of information correctly.	Practical reports.			Statistical analysis reports.
	Use the Internet, MS Word, MS Excel, SAS, NCSS.	Practical reports.			Statistical analysis reports.
Ethically, environmentally and socially aware and active	Discuss ethical Research.	Class discussion; Reading tasks.	Semester test and final exam questions.		
Autonomous and collaborative	Begin to develop life-long learning capabilities and to see one's discipline in a wider context.	Reading and writing tasks.			Statistical analysis reports.
Skilled communicators	Present a clear, well-structured statistical report.	Improve statistical consultation skills; Discussions; Practical reports.			Statistical analysis reports.
Interpersonal flexibility and confidence to engage across difference	Work productively in co-operative learning groups.	Group discussions.			

Semester mark: **70%** of the average of the **test marks**
15% of the average of the **tutorials marks**
15% of the average of the **practical tests**

Final mark: **50%** of the **semester mark**
50% of the **exam mark**

Tests
The tests count 70% towards your semester mark. Test dates are indicated in the module schedule and will take place on Wednesdays in the lunch time lecture period in N22. See academic discipline below for other arrangements.

Tutorials
There is a tutorial class scheduled weekly. It is the responsibility of every student to do the out-of-class **tutorial preparation** that consists of exercises listed in the module schedule and to **review the theory and examples covered in class**. The aim of weekly tutorials is to ensure that you work consistently and stay prepared. The tutorials count 15% towards the semester mark. Each student are also required to work through the self-study exercises.

Practical (Assignments & Tests)
Practical assignments consist of MS EXCEL data work and working in SAS. Assignments will be based on work done in MS EXCEL and tests will be based on work done in SAS. These count 15% towards the final mark.

Final Exam
The final exam consists of all work covered during the semester.

Feedback on Assessment
Feedback on tutorials and tests will take place one week after submission. It is important that you collect your script as soon as it becomes available.

Penalties for Late Submission of Tutorials and Assignments
These penalties will be decided by the Faculty Office.

Special Consideration and Additional Assessments
Only one sick test is scheduled for this module. It will take place on **Monday 11 May 2020 in N22** in the morning lecture. This test will cover chapters 3 to 5. Please take note that this is a sick test and only students that have handed in a sick certificate or appropriate official documents for missing any of the class tests, will be allowed into the venue. **No students will be allowed to use the sick test as a make-up test to improve their semester marks**.

Consult the General Calendar, rule A.5.2.8 for special assessments and A.5.2.16 for resubmission of assessment exercises.

7. Evaluation of the Teaching and Learning

At the end of each term students will be given the opportunity to complete an anonymous pen-and-paper questionnaire and to make some comments about the module. The comments will be summarized to improve presentation of the module in future.

8. Website

Please consult the department website regularly for any additional notes and/or exercises that the lecturer may upload for you to download and work through in your own time. The website is:
https://spsdep.wixsite.com/uwcstatspop/undergraduate-2nd-year

Please consult the Statistical Association of South Africa's (SASA's) website regularly for any 2nd year competitions, scholarships, bursaries. The website is:
http://www.sastat.org.za/

Part C. General Information

9. Academic Honesty and Discipline

Please take the time to read section 3.5 in the General Calendar Part 1. It deals with matters of plagiarism and academic dishonesty.

Students are requested to be on time for lectures, tutorials, tests and practical's.

Make an appointment with the module coordinator for any other queries.

FACULTY OF NATURAL SCIENCES

Department of Statistics & Population Studies

STA 211 – Chapter 2

(Wackerley Mendenhall & Scheaffer, 7th Edition, 2008)

Presentation: R Jacobs

Work through 13 definitions & 9 theorems with Examples.

UNIVERSITY of the WESTERN CAPE

A place of quality, a place to grow, from hope to action through knowledge

Chapter 2 – Probability

Objectives (Revision)

2.1 Introduction

2.2 Probability & Inference

2.3 A Review of Set Notation

2.4 A Probabilistic Model for an Experiment: Discrete Case

2.5 Calculating the Probability of an Event: The Sample-Point Method

2.6 Tools for Counting Sample Points

UWC

Chapter 2 – Probability

Objectives

Revision

2.7 Conditional Probability & the Independence of Events

2.8 Two Laws of Probability

2.9 Calculating the Probability of an Event: The Event-Composition Method

2.10 The **Law of Total Probability** and **Bayes' Rule**

2.11 Numerical Events & Random Variables

2.12 Random Sampling

2.13 Summary

UWC

2.1 Introduction

- Generally **probability is a measure of one's belief** in the occurrence of a future event.

  - **Understand-context**
  - **How-measure**
  - **Assists-making inference**

- **Probability is necessary** when observations generated in the fields of sociology, biology etc. cannot be predicted with certainty. **Examples-** random events such as: **blood pressure of a person** or **when a bridge will collapse**
  - The **relative frequency (rf)** with which such **random events** occur is **stable in a long series of trials**.
  - Such events are called ***random*** or ***stochastic*** events

Probability ~ Long run proportion

2.1 Introduction

- **Example**
  - Its **impossible to predict the occurrence of heads** on a **single toss** of a balanced coin.
  - With a fair measure of confidence the **fraction of heads** in a **long series of trials** would be **very near 0.5**.
- **rf** concept of probability → **intuitively meaningful**
- We **accept a probability interpretation based on rf** (there are many others) as a meaningful measure of our belief in the occurrence of an event.

2.2 Probability & Inference

What is the link that probability provides between observation & inference?

- Gambler wants to **make an inference** (conclusion based on evidence) concerning the **balance of a die**.
  - **Conceptual population**: Set of numbers generated if the die were rolled over & over again (infinitely).
  - If the **die were balanced**, 1/6th - 1's, 1/6th - 2's, etc.

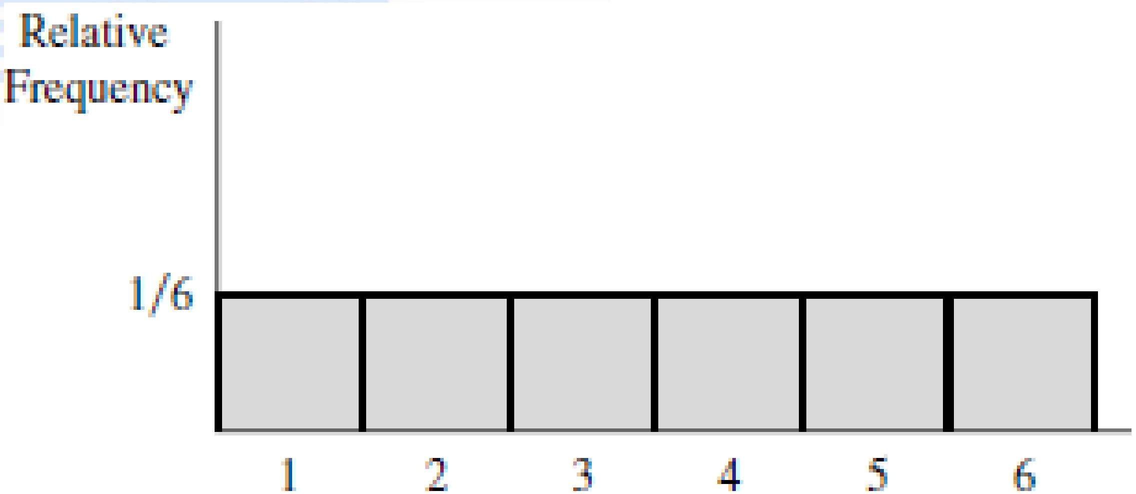


FIGURE 2.1 Frequency distribution for the population generated by a balanced die

<u>2.2 Probability & Inference</u>

- Gambler **hypothesises** - the **die is balanced** & seeks **observations from nature** to contradict the theory, if false.
  - If a **sample** of **10 tosses** (from population) by **rolling the die 10 times** results in **all 1's**.
  - Gambler <u>**concludes**</u> that his <u>**hypothesis is not in agreement with nature**</u> & hence the **die is not balanced**.

The reasoning employed by the gambler <u>identifies the role that probability plays</u> in making inferences.

- Gambler **rejects** his **hypothesis** <u>**not**</u> because it is <u>**impossible to throw 10~1's in 10~tosses**</u> of a balanced die but because it is <u>**highly improbable**</u>.

UWC

2.2 Probability & Inference

- **Evaluation** of the probability most likely **subjective**.
- Gambler may **not have known** how to **calculate** the **probability**, but he had an **intuitive feeling** that this **event was highly unlikely if the die were balanced**. His decision was based on the probability of the observed sample.

- We need a **theory of probability** that will:
  - provide a rigorous **method for finding a number** (a probability) that will **agree with the actual rf** of occurrence of an event in a **long series of trials**.
  - permit us to **calculate the probability of observing specified outcomes,** assuming that our hypothesized model is correct.

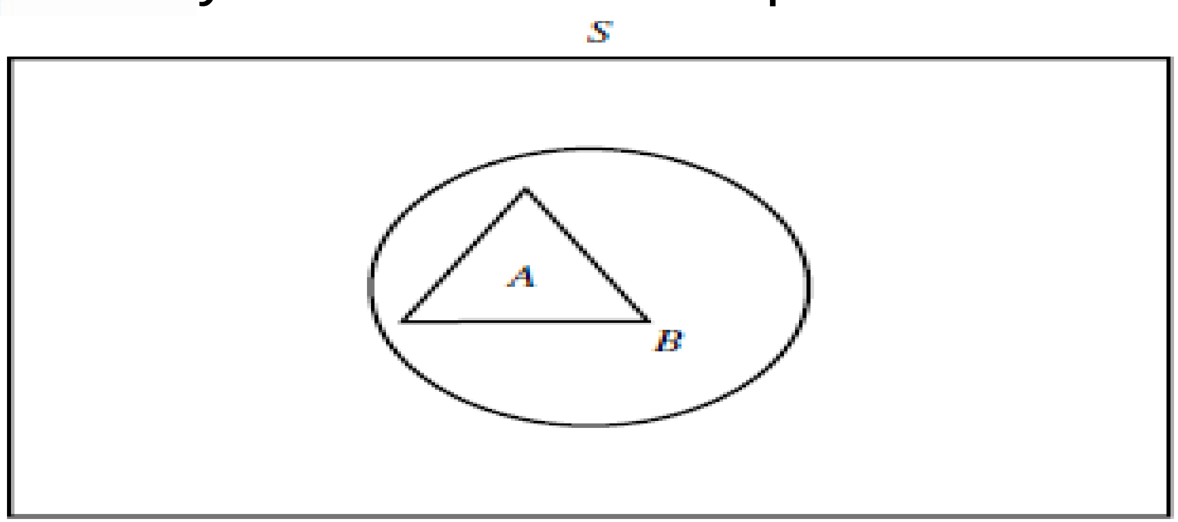
UWC

2.3 A Review of Set Notation

Basic concepts

- **Capital letters**, *A,B...* denote **sets of points**
- **Elements in set** *A;* a_1, a_2, & a_3 : $A = \{a_1, a_2, a_3\}$
- **S – Universal Set (Sample Space)**: A**ll possible outcomes** of a statistical experiment
- **Subset:** For any 2 sets *A* & *B* → *A* is a *subset* of *B* ($A \subset B$), if **every point in *A*** is also **in B**
- ***Null/empty set, Ø:*** Set with no points (Ø ⊂ every set)
- **Venn diagrams :** Portrays sets & relationships between sets.

FIGURE 2.2
Venn diagram for $A \subset B$



2.3 A Review of Set Notation

UNION of A & B: ***A∪B*** - set of all points in *A* **OR** *B* **OR** both. Union of *A & B* contains all points that are in @ least 1of the sets.

INTERSECTION of A & B: **A∩B** / AB - set of all points in both A **AND** B .

FIGURE 2.3
Venn diagram for $A \cup B$

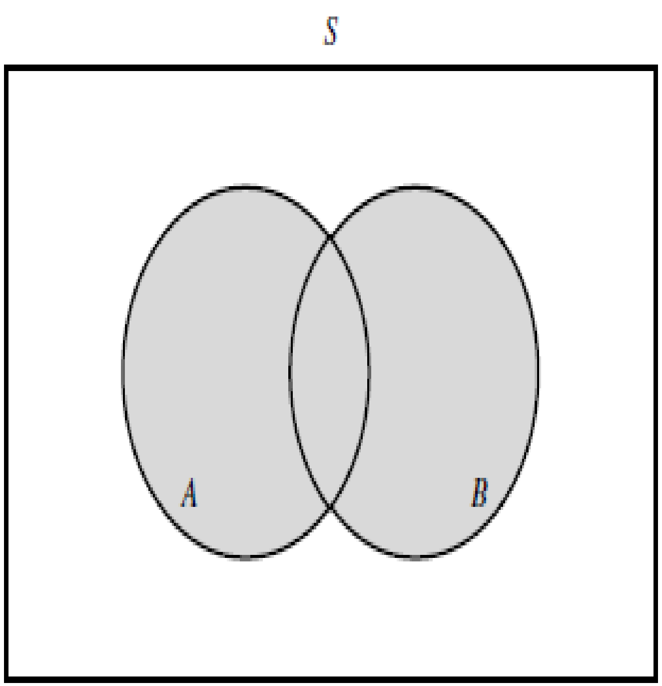
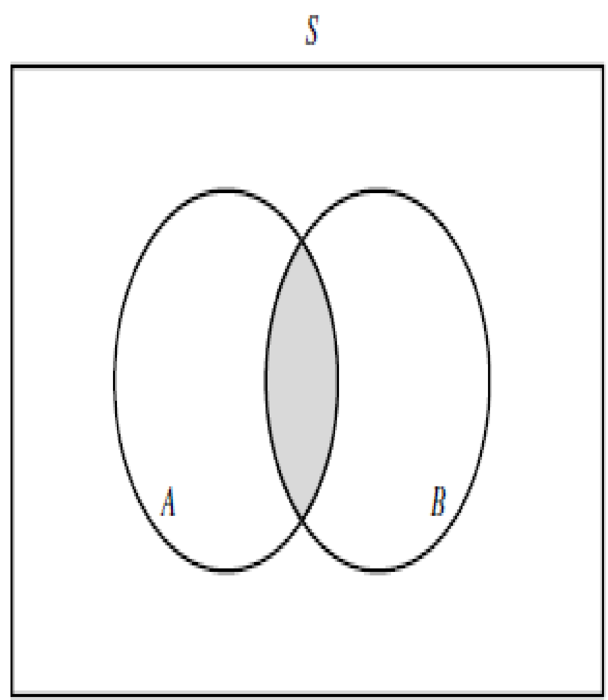


FIGURE 2.4
Venn diagram for AB



2.3 A Review of Set Notation

FIGURE 2.5
Venn diagram for $\overline{A}$

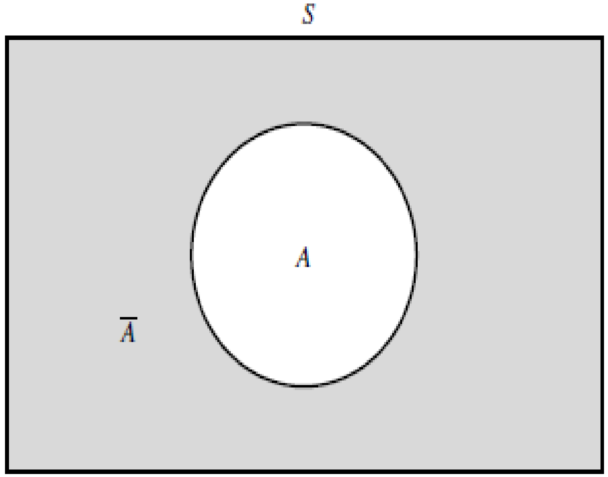
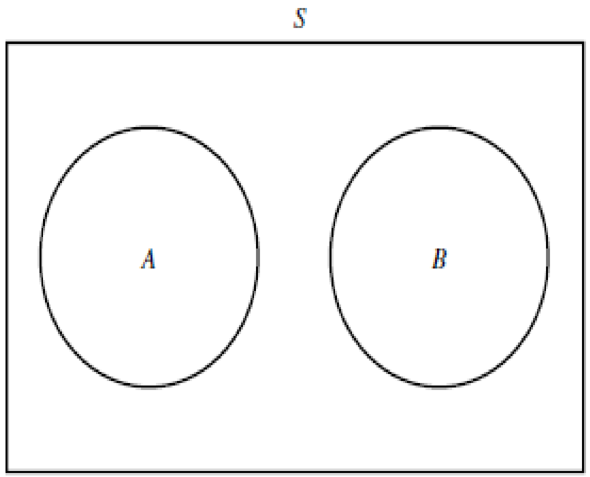


FIGURE 2.6
Venn diagram for mutually exclusive sets A and B



Complement of A ($\overline{A}$)

- If $A \subset S$, $\overline{A}$ is the set of points that are in S but not in A.
- **Note:** $A \cup \overline{A} = S$.

Disjoint/mutually exclusive:

$A \cap B = \emptyset$. Sets have no points in common. *A&B* mutually exclusive.

Exercise: **S**ingle die toss ***S*={1, 2, 3, 4, 5, 6}** Let ***A*={1,2}, *B*={1,3}, C={2,4,6}**

$A \cup B$ = **?**, $A \cap B$ =?, $\overline{A}$= **?**, Are B & *C* and *A* & *C* are mutually exclusive **?**

2.3 A Review of Set Notation

Example: ***S*={1,2,3,4,5,6}** Let ***A*={1,2}, *B*={1,3}, C={2,4,6}**

$A \cup B = \{1, 2, 3\}$, $A \cap B = \{1\}$, $\overline{A} = \{3,4,5,6\}$, ***B & C* are** mutually exclusive ***A & C* are not.**

Distributive laws:

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$$

$$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$$

De Morgan's laws:

$$\overline{(A \cap B)} = \overline{A} \cup \overline{B}$$

$$\overline{(A \cup B)} = \overline{A} \cap \overline{B}$$

2.4 A Probabilistic Model for an Experiment: The Discrete Case

Definition 2.1: An **experiment** is the process by which an observation is made.

Examples: coin & die tossing, measuring the an individual's IQ

Events (denoted by capital letters)**:** Outcomes of an experiment.

Example: Counting the no. bacteria in a portion of food, events..

- *A*: Exactly 110 bacteria are present.
- *B*: More than 200 bacteria are present.
- *C*: The number of bacteria present is between 100 & 300

2.4 A Probabilistic Model for an Experiment: The Discrete Case

Definition 2.2:

– A ***simple event* ~** event that cannot be decomposed

– Has only **1 sample point**

– The letter **E** with a subscript denotes a simple event

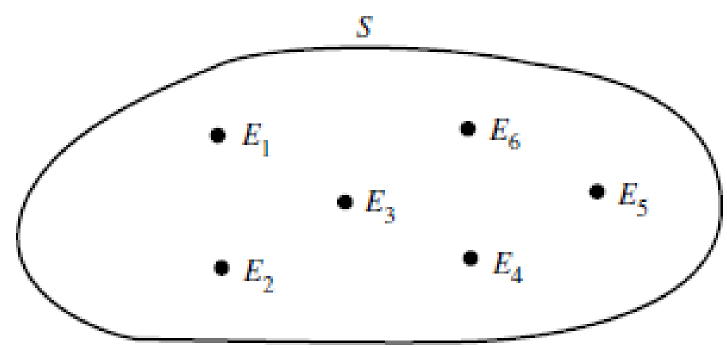
Definition 2.3:

– The ***sample space* ~** set consisting of **all possible sample points**

– Denoted by **S**

– Can be **finite** or **infinite**

FIGURE 2.7 Venn diagram for the sample space associated with the die-tossing experiment



2.4 A Probabilistic Model for an Experiment: The Discrete Case

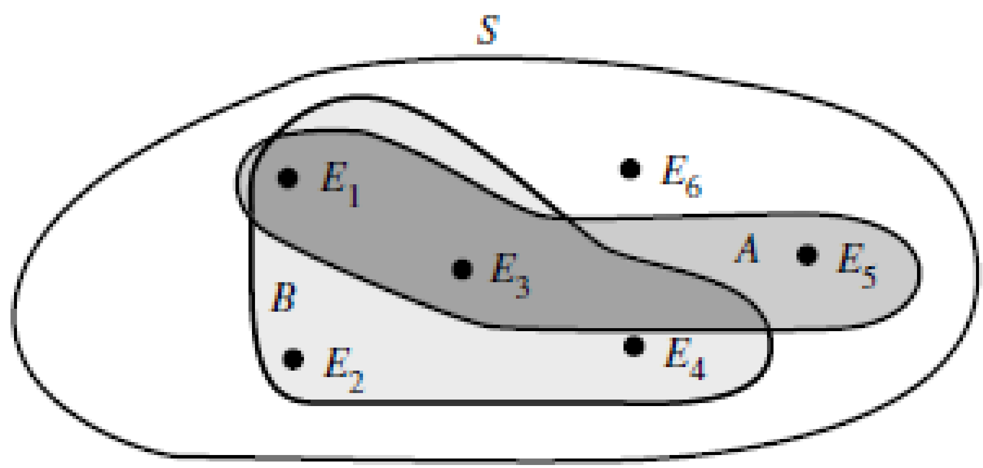
Definition 2.4:

A ***discrete sample space*** is one that contains either a finite or a countable number of distinct sample points.

Definition 2.5:

An ***event in a discrete sample*** space *S* is a collection of sample points - that is, any subset of *S*

FIGURE 2.8
Venn diagram for the die-tossing experiment



2.4 A Probabilistic Model for an Experiment: The Discrete Case

- **Probabilistic model** for an experiment with a **discrete S** can be constructed by ***assigning a numerical probability*** **to each <u>simple</u> E in the S**.

- We will **select this number**, a **measure of our belief** in the **event's occurrence** on a **single repetition** of the experiment, in such a way that it will be ***consistent with*** the ***r.f. concept of probability***.

2.4 A Probabilistic Model for an Experiment: The Discrete Case

NB: rf concept of probability - 3 conditions must hold:

1. The **rf** of occurrence of any event must be **≥ 0**, No negative rf.

2. The **rf of** the whole **S** must **be 1**. Every possible outcome of the experiment is a point in S→S must occur every time the experiment is performed.

3. If **2 events** are **mutually exclusive**, the **rf of their ∪** is the **∑ their respective rf**.

UWC

2.4 A Probabilistic Model for an Experiment: The Discrete Case

Definition 2.6:

Suppose S is a sample space associated with an experiment. To every event A in S (A is a subset of S), we assign a number, $P(A)$, called the *probability* of A, so that the following axioms hold:

Axiom 1: $P(A) \geq 0$. √

Axiom 2: $P(S) = 1$. √

Axiom 3: If $A_1, A_2, A_3, \ldots$ form a sequence of pairwise mutually exclusive events in S (that is, $A_i \cap A_j = \emptyset$ if $i \neq j$), then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_{i=1}^{\infty} P(A_i).$$ √

Review Example 2.1, p31

UWC

2.5 Calculating the Probability of an Event: The Sample-Point Method

The sample-point method is outlined in Section 2.4. The following steps are used to find the probability of an event:

1. Define the experiment and clearly determine how to describe one simple event.
2. List the simple events associated with the experiment and test each to make certain that it cannot be decomposed. This defines the sample space S.
3. Assign reasonable probabilities to the sample points in S, making certain that $P(E_i) \geq 0$ and $\sum P(E_i) = 1$.
4. Define the event of interest, A, as a specific collection of sample points. (A sample point is in A if A occurs when the sample point occurs. Test *all* sample points in S to identify those in A.)
5. Find $P(A)$ by summing the probabilities of the sample points in A.

UWC

2.5 Calculating the Probability of an Event: The Sample-Point Method

Example 2.2:

Consider the problem of selecting two applicants for a job out of a group of five and imagine that the applicants vary in competence, 1 being the best, 2 second best, and so on, for 3, 4, and 5. These ratings are of course unknown to the employer. Define two events A and B as:

A: The employer selects the best and one of the two poorest applicants (applicants 1 and 4 or 1 and 5).

B: The employer selects at least one of the two best.

Find the probabilities of these events.

Follow the 5 steps

UWC

Example 2.2 Solution:

1. The experiment involves randomly selecting two applicants out of five. Denote the selection of applicants 3 and 5 by {3, 5}.
2. The ten simple events, with $\{i, j\}$ denoting the selection of applicants i and j, are

$$
\begin{array}{llll}
E_1: \{1, 2\}, & E_5: \{2, 3\}, & E_8: \{3, 4\}, & E_{10}: \{4, 5\}. \\
E_2: \{1, 3\}, & E_6: \{2, 4\}, & E_9: \{3, 5\}, & \\
E_3: \{1, 4\}, & E_7: \{2, 5\}, & & \\
E_4: \{1, 5\}, & & &
\end{array}
$$

3. A random selection of two out of five gives each pair an equal chance for selection. Hence, we will assign each sample point the probability 1/10. That is,

$$P(E_i) = 1/10 = .1, \qquad i = 1, 2, \ldots, 10.$$

P(B)?

4. Checking the sample points, we see that B occurs whenever E_1, E_2, E_3, E_4, E_5, E_6, or E_7 occurs. Hence, these sample points are included in B.

Example 2.2 Solution:

5. Finally, $P(B)$ is equal to the sum of the probabilities of the sample points in B, or

$$P(B) = \sum_{i=1}^{7} P(E_i) = \sum_{i=1}^{7} .1 = .7.$$

Similarly, we see that event $A = E_3 \cup E_4$ and that $P(A) = .1 + .1 = .2.$ ■

Sample-point method for solving a probability problem is:

- direct & **powerful**,
- not resistant to **human error** (example; incorrectly diagnosing the nature of a simple event or failing to list all the sample points in *S*).

2.6 Tools for Counting Sample Points

Using **Combinatorial Analysis** to find the **P(event)**

- If **E⊂S is large** & manual enumeration of every sample point is tedious/impossible
- Sometimes **counting the no. of points** in **S & E** may be the **only way** to **calculate the probability of an E** .

- **S** contains **N** equiprobable sample points
- **Event *A*** contains n_a sample points, $P(A)=n_a/N$

2.6 Tools for Counting Sample Points

Theorem 2.1: (*mn rule*)

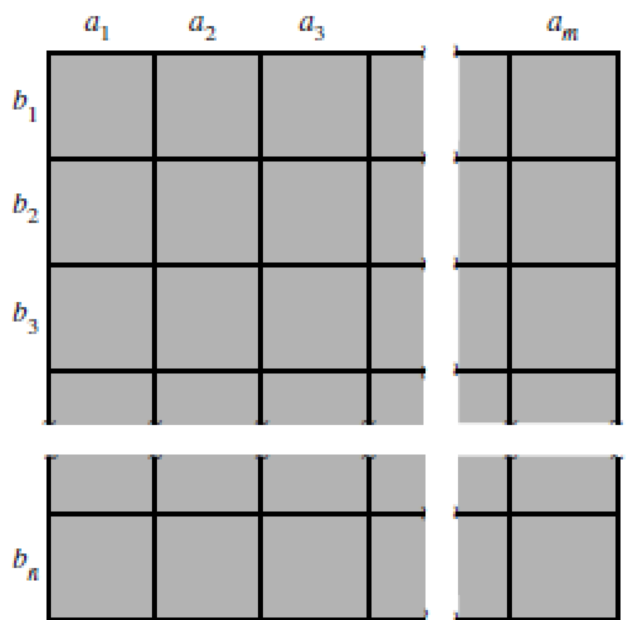
With m elements $a_1, a_2, \ldots, a_m$ and n elements $b_1, b_2, \ldots, b_n$, it is possible to form $mn = m \times n$ pairs containing one element from each group.

Proof

Verification of the theorem can be seen by observing the rectangular table in Figure 2.9. There is one square in the table for each a_i, b_j pair and hence a total of $m \times n$ squares.

The *mn* rule **can be extended** to any number of sets.

FIGURE 2.9
Table indicating the number of pairs (a_i, b_j)



2.6 Tools for Counting Sample Points

Example 2.6:

A balanced coin is tossed 3 times

Refer to the coin-tossing experiment in Example 2.3. We found for this example that the total number of sample points was eight. Use the extension of the mn rule to confirm this result.

$E_1: HHH,\quad E_3: HTH,\quad E_5: HTT,\quad E_7: TTH,$
$E_2: HHT,\quad E_4: THH,\quad E_6: THT,\quad E_8: TTT.$

Solution:

Each sample point in S was identified by a sequence of three letters, where each position in the sequence contained one of two letters, an H or a T. The problem therefore involves the formation of triples, with an element (an H or a T) from each of three sets. For this example the sets are identical and all contain two elements (H and T). Thus, the number of elements in each set is $m = n = p = 2$, and the total number of triples that can be formed is $mnp = (2)^3 = 8$. ■

UWC

2.6 Tools for Counting Sample Points

Definition 2.7:

Order is important

An ordered arrangement of r distinct objects is called a *permutation*. The number of ways of ordering n distinct objects taken r at a time will be designated by the symbol P_r^n.

Theorem 2.2: Permutations

Why (n-r+1)?

r ≤ n?

$$P_r^n = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}.$$

How would we go about proving this theorem?

Factorial (n!): Product of all the natural numbers from n to 1.

Exercise: 3! = **?**

UWC

2.6 Tools for Counting Sample Points

Theorem 2.2 Proof:

We are concerned with the number of ways of filling r positions with n distinct objects. Applying the extension of the mn rule, we see that the first object can be chosen in one of n ways. After the first is chosen, the second can be chosen in $(n-1)$ ways, the third in $(n-2)$, and the rth in $(n-r+1)$ ways. Hence, the total number of distinct arrangements is

$$P_r^n = n(n-1)(n-2)\cdots(n-r+1).$$

Divide by the remainder (interest in the r^{th} object)

Expressed in terms of factorials,

$$P_r^n = n(n-1)(n-2)\cdots(n-r+1)\frac{(n-r)!}{(n-r)!} = \frac{n!}{(n-r)!}$$

where $n! = n(n-1)\cdots(2)(1)$ and $0! = 1$.

NB:

n=n-(1-1), n-1=n–(2-1),….,**n-(r-1)=(n-r+1), n-(r+1-1)=(n-r),** n-(r+2-1)=(n-r-1)...

1st position, 2nd position,……rth position, r+1th position, r+2th position…

n! = n*(n-1)…*(n-r+1)*(n-r)(n-r-1)…3*2*1

2.6 Tools for Counting Sample Points

Example 2.9:

Suppose that an assembly operation in a manufacturing plant involves four steps, which can be performed in any sequence. If the manufacturer wishes to compare the assembly time for each of the sequences, how many different sequences will be involved in the experiment?

Solution:

The total number of sequences equals the number of ways of arranging the $n = 4$ steps taken $r = 4$ at a time, or

$$P_4^4 = \frac{4!}{(4-4)!} = \frac{4!}{0!} = 24. \quad \blacksquare$$

UWC

How many sequences if order was not important?

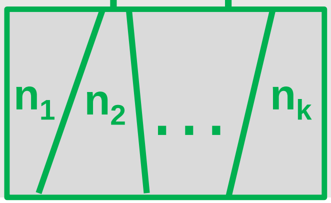
2.6 Tools for Counting Sample Points

Theorem 2.3: Partitions

Dividing

The number of ways of partitioning n distinct objects into k distinct groups containing $n_1, n_2, \ldots, n_k$ objects, respectively, where each object appears in exactly one group and $\sum_{i=1}^{k} n_i = n$, is

Sample Space



$$N = \binom{n}{n_1 \; n_2 \; \cdots \; n_k} = \frac{n!}{n_1! \, n_2! \cdots n_k!}.$$

$N = \binom{n}{n_1 \ldots n_k}$

How would we go about proving this theorem?

Theorem 2.3 Proof:

N is the number of distinct arrangements of n objects in a row for a case in which rearrangement of the objects within a group does not count. For example, the letters a to l are arranged in three groups, where $n_1 = 3, n_2 = 4$, and $n_3 = 5$: $abc|defg|hijkl$ is one such arrangement.

OR bca|gdef|lhijk etc.

2.6 Tools for Counting Sample Points

$$P_r^n = n(n-1)(n-2)\cdots(n-r+1) = \frac{n!}{(n-r)!}.$$

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

Theorem 2.3 Proof:

The number of distinct arrangements of the n objects, assuming all objects are distinct, is $P_n^n = n!$ (from Theorem 2.2). Then P_n^n equals the number of ways of partitioning the n objects into k groups (ignoring order within groups) multiplied by the number of ways of ordering the $n_1, n_2, \ldots, n_k$ elements within each group. This application of the extended mn rule gives

$$P_n^n = (N) \cdot (n_1!\, n_2!\, n_3! \cdots n_k!),$$

where $n_i!$ is the number of distinct arrangements of the n_i objects in group i.

Solving for N, we have

Equality principal

$$N = \frac{n!}{n_1!\, n_2! \cdots n_k!} \equiv \binom{n}{n_1\; n_2\; \cdots\; n_k}.$$

Do Example 2.10,p45

2.6 Tools for Counting Sample Points Sampling

Definition 2.8:

Order is NOT important

The number of *combinations* of n objects taken r at a time is the number of subsets, each of size r, that can be formed from the n objects. This number will be denoted by C_r^n or $\binom{n}{r}$.

Example 2.11 + Solution:

Find the number of ways of selecting two applicants out of five

$$\binom{5}{2} = \frac{5!}{2!3!} = 10.$$

Theorem 2.4:

The number of unordered subsets of size r chosen (without replacement) from n available objects is

$$\binom{n}{r} = C_r^n = \frac{P_r^n}{r!} = \frac{n!}{r!(n-r)!}.$$

Where would we start to prove this?

2.6 Tools for Counting Sample Points

$$N = \binom{n}{n_1 \; n_2 \; \cdots \; n_k} = \frac{n!}{n_1! \, n_2! \cdots n_k!}.$$

Theorem 2.4 Proof:

The selection of r objects from a total of n is equivalent to partitioning the n objects into $k = 2$ groups, the r selected, and the $(n - r)$ remaining. This is a special case of the general partitioning problem dealt with in Theorem 2.3. In the present case, $k = 2$, $n_1 = r$, and $n_2 = (n - r)$ and, therefore,

Binomial Coefficients

$$\binom{n}{r} = C_r^n = \binom{n}{r \quad n-r} = \frac{n!}{r!(n-r)!}.$$

Example 2.12,p46

Let A denote the event that exactly one of the two best applicants appears in a selection of two out of five. Find the number of sample points in A and $P(A)$.

UWC

2.6 Tools for Counting Sample Points

Example 2.12 Solution:

Let n_a denote the number of sample points in A. Then n_a equals the number of ways of selecting one of the two best (call this number m) times the number of ways of selecting one of the three low-ranking applicants (call this number n). Then $m = \binom{2}{1}$, $n = \binom{3}{1}$, and applying the mn rule,

$$n_a = \binom{2}{1} \cdot \binom{3}{1} = \frac{2!}{1!1!} \cdot \frac{3!}{1!2!} = 6.$$

In Example 2.11 we found the total number of sample points in S to be $N = 10$. If each selection is equiprobable, $P(E_i) = 1/10 = .1$, $i = 1, 2, \ldots, 10$, and

$$P(A) = \sum_{E_i \subset A} P(E_i) = \sum_{E_i \subset A} (.1) = n_a(.1) = 6(.1) = .6.$$

$$\binom{5}{2} = \frac{5!}{2!3!} = 10.$$

2.7 Conditional Probability & the Independence of Events

Probability dependent on occurrence of prior events

Definition 2.9:

The *conditional probability of an event A*, given that an event B has occurred, is equal to

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

provided $P(B) > 0$. [The symbol $P(A|B)$ is read "probability of A given B."]

Do Example 2.14,p53

2.7 Conditional Probability & the Independence of Events

Probability of the occurrence of an event is **unaffected** by the **occurrence/non-occurrence** of another event.

Definition 2.10

Two events A and B are said to be *independent* if any one of the following holds:

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

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

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

Otherwise, the events are said to be *dependent*.

Do Example 2.16,p54

2.8 Two Laws of Probability

Theorem 2.5 (Multiplication→∩ of E's)

The Multiplicative Law of Probability The probability of the intersection of two events A and B is

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

If A and B are independent, then

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

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

Proof

The multiplicative law follows directly from Definition 2.9, the definition of conditional probability.

Extension of the multiplicative law:

$P(A \cap B \cap C)$

$= P[(A \cap B) \cap C] = P(A \cap B)P(C|A \cap B) = P(A)P(B|A)P(C|A \cap B)$

$P(A_1 \cap A_2 \cap A_3 \cap \cdot \cdot \cdot \cap A_k)$

$= P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdot \cdot \cdot P(A_k|A_1 \cap A_2 \cap \cdot \cdot \cdot \cap A_{k-1})$

2.8 Two Laws of Probability

Theorem 2.6 (Addition → U of E)

The Additive Law of Probability The probability of the union of two events A and B is

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

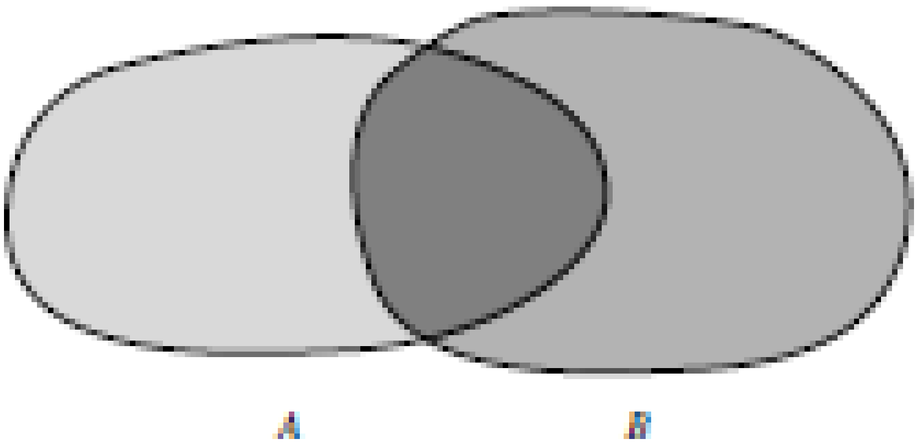
If A and B are mutually exclusive events, $P(A \cap B) = 0$ and

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

Theorem 2.6 Proof:

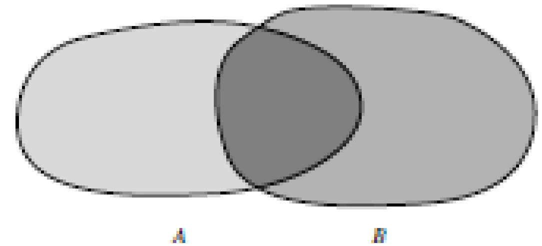
How can you use the Venn to prove this Theorem?

FIGURE 2.10
Venn diagram for the union of A and B



2.8 Two Laws of Probability

Theorem 2.6 Proof:



Notice that $A \cup B = A \cup (\overline{A} \cap B)$, where A and $(\overline{A} \cap B)$ are mutually exclusive events. Further, $B = (\overline{A} \cap B) \cup (A \cap B)$, where $(A \cap \overline{B})$ and $(A \cap B)$ are mutually exclusive events. Then, by Axiom 3,

Axiom 3: If $A_1, A_2, A_3, \ldots$ form a sequence of pairwise mutually exclusive events in S (that is, $A_i \cap A_j = \emptyset$ if $i \neq j$), then

$$P(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_{i=1}^{\infty} P(A_i).$$

$$P(A \cup B) = P(A) + P(\overline{A} \cap B) \quad \text{and} \quad P(B) = P(\overline{A} \cap B) + P(A \cap B).$$

The equality given on the right implies that $P(\overline{A} \cap B) = P(B) - P(A \cap B)$. Substituting this expression for $P(\overline{A} \cap B)$ into the expression for $P(A \cup B)$ given in the left-hand equation of the preceding pair, we obtain the desired result:

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

2.8 Two Laws of Probability

Union of 3 events:

$P(A \cup B \cup C)$

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

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

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

NB: $(A \cap B) \cap (A \cap C) = A \cap B \cap C$.

Theorem 2.7: (Compliment Law)

Sometimes it is easier to calculate the compliment than to calculate *P(A)*.

Proof

If A is an event, then

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

Observe that $S = A \cup \overline{A}$. Because A and $\overline{A}$ are mutually exclusive events, it follows that $P(S) = P(A) + P(\overline{A})$. Therefore, $P(A) + P(\overline{A}) = 1$ and the result follows.

2.9 Calculating the Probability of an Event: The Event-Composition Method

Event compositions involves **unions** and/or **intersections**.

Example 2.17:

Of the voters in a city, 40% are Republicans and 60% are Democrats. Among the Republicans 70% are in favor of a bond issue, whereas 80% of the Democrats favor the issue. If a voter is selected at random in the city, what is the probability that he or she will favor the bond issue?

Exercise: Draw the Venn diagram.

2.9 Calculating the Probability of an Event: The Event-Composition Method

Example 2.17 – Solution:

Let F denote the event "favor the bond issue," R the event "a Republican is selected," and D the event "a Democrat is selected." Then $P(R) = .4$, $P(D) = .6$, $P(F|R) = .7$, and $P(F|D) = .8$. Notice that

$$P(F) = P[(F \cap R) \cup (F \cap D)] = P(F \cap R) + P(F \cap D)$$

because $(F \cap R)$ and $(F \cap D)$ are mutually exclusive events.

It follows that

$$P(F \cap R) = P(F|R)P(R) = (.7)(.4) = .28,$$

$$P(F \cap D) = P(F|D)P(D) = (.8)(.6) = .48.$$

$$P(F) = .28 + .48 = .76.$$

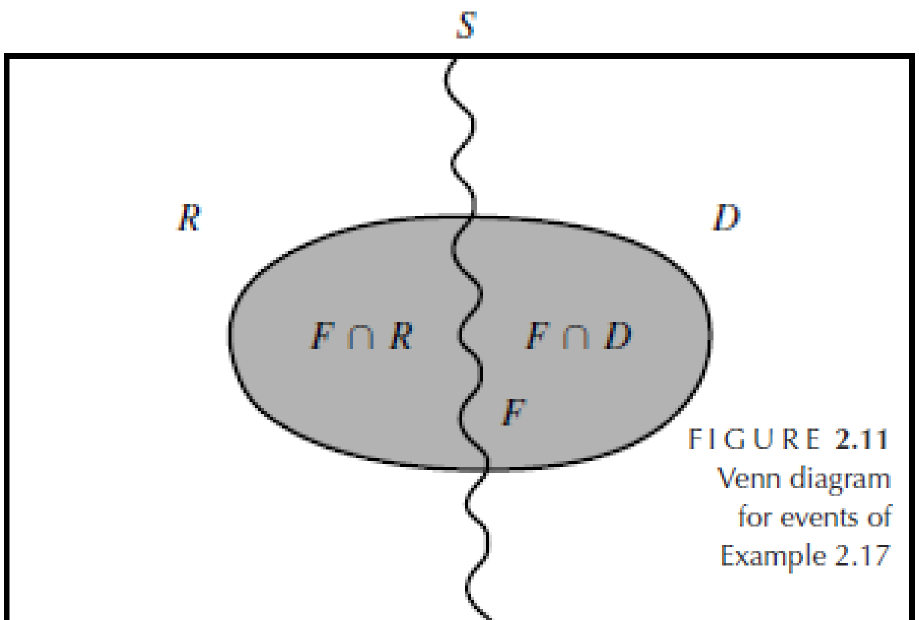


FIGURE 2.11 Venn diagram for events of Example 2.17

Do Example 2.18, pg63

2.9 Calculating the Probability of an Event: The Event-Composition Method

A summary of the steps used in the event-composition method follows:

1. Define the experiment.
2. Visualize the nature of the sample points. Identify a few to clarify your thinking.
3. Write an equation expressing the event of interest—say, A—as a composition of two or more events, using unions, intersections, and/or complements. (Notice that this equates point sets.) Make certain that event A and the event implied by the composition represent the same set of sample points.
4. Apply the additive and multiplicative laws of probability to the compositions obtained in step 3 to find $P(A)$.

Example 2.19:

Two applicants are randomly selected from among five who have applied for a job. Find the probability that exactly one of the two best applicants is selected, event A.

2.9 Calculating the Probability of an Event: The Event-Composition Method

Example 2.19:

Two applicants are randomly selected from among five who have applied for a job. Find the probability that exactly one of the two best applicants is selected, event A.

Example 2.19 Solution Discussion:

Define the following two events:

B: Draw the best and one of the three poorest applicants.
C: Draw the second best and one of the three poorest applicants.

Events B and C are mutually exclusive and $A = B \cup C$. Also, let $D_1 = B_1 \cap B_2$, where

B_1 = Draw the best on the first draw,
B_2 = Draw one of the three poorest applicants on the second draw,

and $D_2 = B_3 \cap B_4$, where

B_3 = Draw one of the three poorest applicants on the first draw,
B_4 = Draw the best on the second draw.

Note that $B = D_1 \cup D_2$.

Example 2.19 Solution:

Similarly, let $G_1 = C_1 \cap C_2$ and $G_2 = C_3 \cap C_4$, where C_1, C_2, C_3, and C_4 are defined like B_1, B_2, B_3, and B_4, with the words *second best* replacing *best*. Notice that D_1 and D_2 and G_1 and G_2 are pairs of mutually exclusive events and that

$$A = B \cup C = (D_1 \cup D_2) \cup (G_1 \cup G_2),$$

$$A = (B_1 \cap B_2) \cup (B_3 \cap B_4) \cup (C_1 \cap C_2) \cup (C_3 \cap C_4).$$

Applying the additive law of probability to these four mutually exclusive events, we have

$$P(A) = P(B_1 \cap B_2) + P(B_3 \cap B_4) + P(C_1 \cap C_2) + P(C_3 \cap C_4).$$

Applying the multiplicative law, we have $P(B_1 \cap B_2) = P(B_1)P(B_2|B_1)$.

The probability of drawing the best on the first draw is $P(B_1) = 1/5$.

Similarly, the probability of drawing one of the three poorest on the second draw, given that the best was drawn on the first selection, is $P(B_2|B_1) = 3/4$.

Then $P(B_1 \cap B_2) = P(B_1)P(B_2|B_1) = (1/5)(3/4) = 3/20.$

Example 2.19 Solution:

The probabilities of all other intersections in $P(A)$, $P(B_3 \cap B_4)$, $P(C_1 \cap C_2)$, and $P(C_3 \cap C_4)$ are obtained in exactly the same manner, and all equal 3/20. Then

$$P(A) = P(B_1 \cap B_2) + P(B_3 \cap B_4) + P(C_1 \cap C_2) + P(C_3 \cap C_4)$$
$$= (3/20) + (3/20) + (3/20) + (3/20) = 3/5.$$

This answer is identical to that obtained in Example 2.12, where $P(A)$ was calculated by using the sample-point approach. ■

Do Examples 2.20 - 2.22 p65-68

UWC

2.10 The Law of Total Probability & Bayes' Rule

Event-composition approach: view *S* as a ∪ of mutually exclusive subsets → **Law of total probability.**

Definition 2.11:

For some positive integer k, let the sets $B_1, B_2, \ldots, B_k$ be such that

1. $S = B_1 \cup B_2 \cup \cdots \cup B_k$.
2. $B_i \cap B_j = \emptyset$, for $i \neq j$.

Then the collection of sets $\{B_1, B_2, \ldots, B_k\}$ is said to be a *partition* of S.

If A is any subset of S and $\{B_1, B_2, \ldots, B_k\}$ is a partition of S, A can be *decomposed* as follows:

$$A = (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k).$$

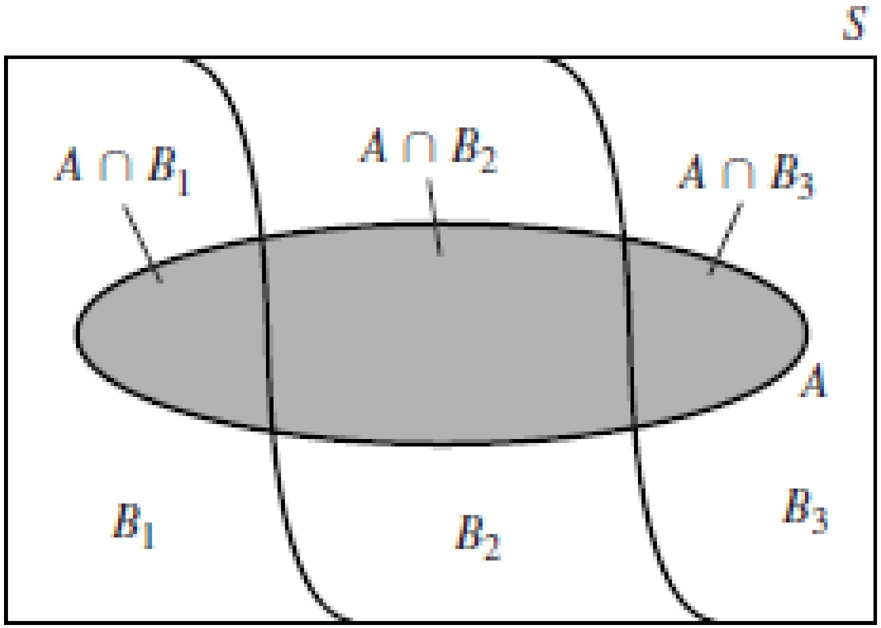
UWC

Family of disjoint events

2.10 The Law of Total Probability & Bayes' Rule

Figure 2.12 illustrates this decomposition for $k = 3$.

FIGURE 2.12
Decomposition of event A



Theorem 2.8:

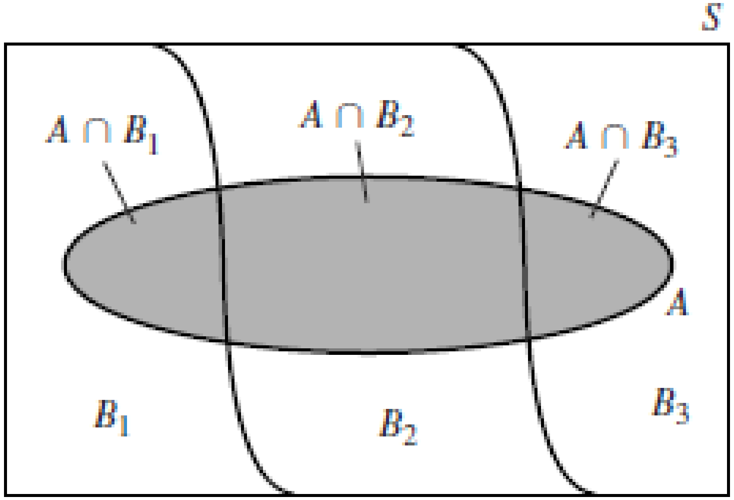
Assume that $\{B_1, B_2, \ldots, B_k\}$ is a partition of S (see Definition 2.11) such that $P(B_i) > 0$, for $i = 1, 2, \ldots, k$. Then for any event A

$$P(A) = \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

Expression in simplest form that will enable us to calculate a probability

2.10 The Law of Total Probability & Bayes' Rule

Theorem 2.8 Proof:



Any subset A of S can be written as

$$A = A \cap S = A \cap (B_1 \cup B_2 \cup \cdots \cup B_k)$$
$$= (A \cap B_1) \cup (A \cap B_2) \cup \cdots \cup (A \cap B_k).$$

Notice that, because $\{B_1, B_2, \cdots, B_k\}$ is a partition of S, if $i \neq j$,

$$(A \cap B_i) \cap (A \cap B_j) = A \cap (B_i \cap B_j) = A \cap \emptyset = \emptyset$$

and that $(A \cap B_i)$ and $(A \cap B_j)$ are mutually exclusive events. Thus,

$$P(A) = P(A \cap B_1) + P(A \cap B_2) + \cdots + P(A \cap B_k)$$
$$= P(A|B_1)P(B_1) + P(A|B_2)P(B_2) + \cdots + P(A|B_k)P(B_k)$$
$$= \sum_{i=1}^{k} P(A|B_i)P(B_i).$$

Dependence

2.10 The Law of Total Probability & Bayes' Rule

Theorem 2.9:

Bayes' Rule Assume that $\{B_1, B_2, \ldots, B_k\}$ is a partition of S (see Definition 2.11) such that $P(B_i) > 0$, for $i = 1, 2, \ldots, k$. Then

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

Theorem 2.9 Proof:

The proof follows directly from the definition of conditional probability and the law of total probability. Note that

$$P(B_j|A) = \frac{P(A \cap B_j)}{P(A)} = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

Law of Multiplication -dependence

Law of total probability

Example 2.23:

An electronic fuse is produced by five production lines in a manufacturing operation. The fuses are costly, are quite reliable, and are shipped to suppliers in 100-unit lots. Because testing is destructive, most buyers of the fuses test only a small number of fuses before deciding to accept or reject lots of incoming fuses.

All five production lines produce fuses at the same rate and normally produce only 2% defective fuses, which are dispersed randomly in the output. Unfortunately, production line 1 suffered mechanical difficulty and produced 5% defectives during the month of March. This situation became known to the manufacturer after the fuses had been shipped. A customer received a lot produced in March and tested three fuses. One failed. What is the probability that the lot was produced on line 1? What is the probability that the lot came from one of the four other lines?

What probabilities are given?

2.10 The Law of Total Probability & Bayes' Rule

Example 2.23 Solution:

Let B denote the event that a fuse was drawn from line 1 and let A denote the event that a fuse was defective. Then it follows directly that

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^{k} P(A|B_i)P(B_i)}.$$

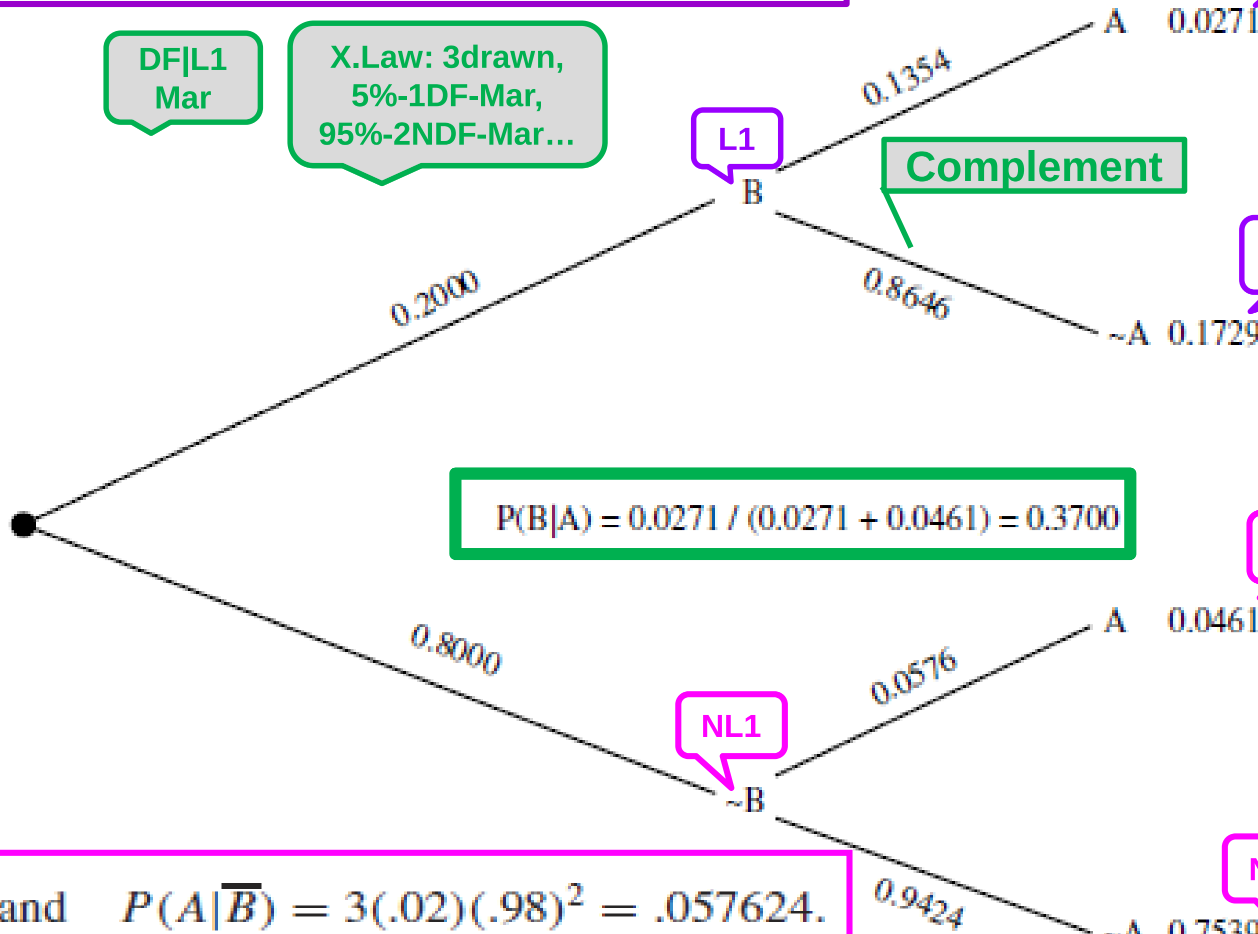
$$P(B) = 0.2 \quad \text{and} \quad P(A|B) = 3(.05)(.95)^2 = .135375.$$

Partition

L1: 1/5

DF|L1 Mar

X.Law: 3drawn, 5%-1DF-Mar, 95%-2NDF-Mar…



P(B|A) = 0.0271 / (0.0271 + 0.0461) = 0.3700

FIGURE 2.13 Tree diagram for calculations in Example 2.23. $\sim A$ and $\sim B$ are alternative notations for $\overline{A}$ and $\overline{B}$, respectively.

$$P(\overline{B}) = 0.8 \quad \text{and} \quad P(A|\overline{B}) = 3(.02)(.98)^2 = .057624.$$

2.10 The Law of Total Probability & Bayes' Rule

$$P(B) = 0.2 \quad \text{and} \quad P(A|B) = 3(.05)(.95)^2 = .135375.$$

Similarly,

$$P(\overline{B}) = 0.8 \quad \text{and} \quad P(A|\overline{B}) = 3(.02)(.98)^2 = .057624.$$

Note that these conditional probabilities were very easy to calculate. Using the law of total probability,

$$P(A) = P(A|B)P(B) + P(A|\overline{B})P(\overline{B})$$
$$= (.135375)(.2) + (.057624)(.8) = .0731742.$$

Finally,

Bayes Rule

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)} = \frac{(.135375)(.2)}{.0731742} = .37,$$

and

$$P(\overline{B}|A) = 1 - P(B|A) = 1 - .37 = .63.$$

Figure 2.13, obtained using the applet *Bayes' Rule as a Tree*, illustrates the various steps in the computation of $P(B|A)$. ■

2.11 Numerical Events & Random Variables(RV)

Numerical Events-Example: 10 of 10 treated patients survive an illness; sales next year will be R5m etc.

RV Y: A variable measured in an experiment, varies depending on the outcome of the experiment.

- Assign to each point in the S **a** $\mathbb{R}$**-value** for RV Y
- **Y** value will **vary from 1 sample point to another** (some may be assigned the same numerical value).

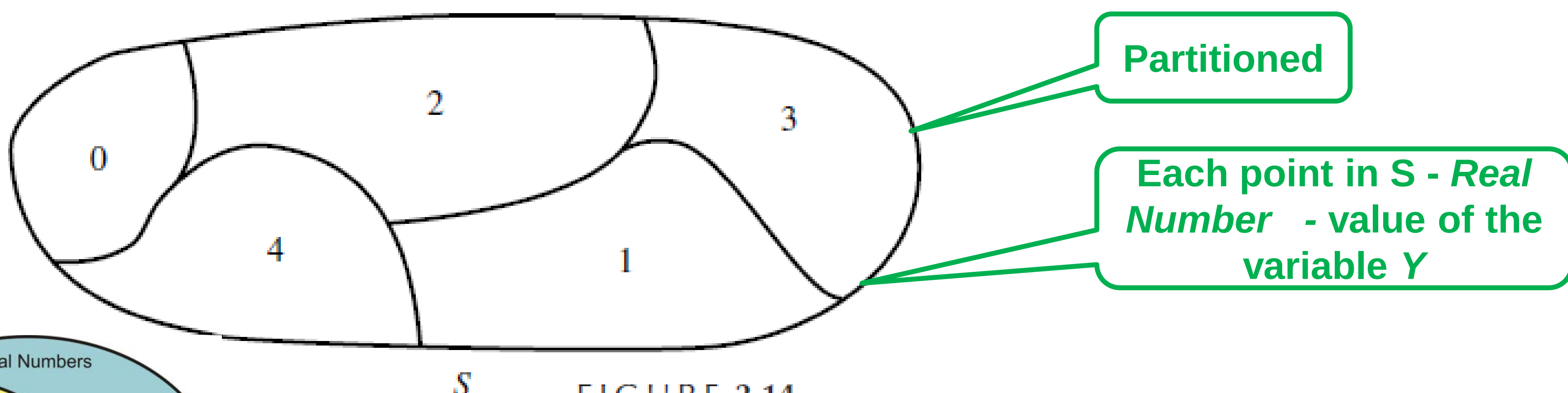
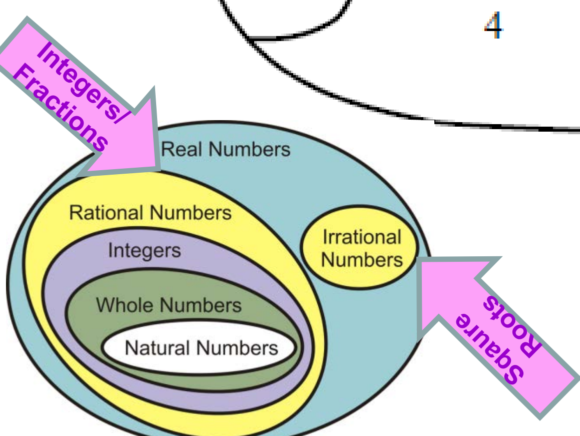
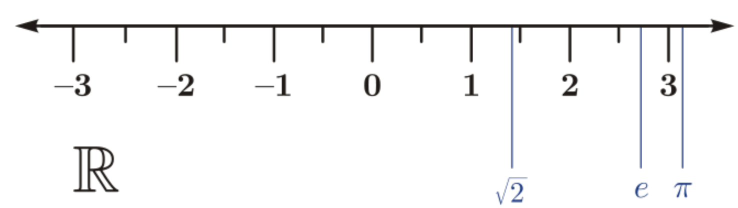


FIGURE 2.14
Partitioning S into subsets that define the events $Y = 0, 1, 2, 3,$ and 4





2.11 Numerical Events & RV's

- A function of the sample points in S & all sample points where **Y = a**, is the **numerical event assigned the number a**.
- **S can be partitioned** into subsets so that points within a subset are all assigned the same value of Y .

Definition 2.12:

A *random variable* is a real-valued function for which the domain is a sample space.

Example 2.24:

Define an experiment as tossing two coins and observing the results. Let Y equal the number of heads obtained. Identify the sample points in S, assign a value of Y to each sample point, and identify the sample points associated with each value of the random variable Y.

SOLUTION?

Example 2.25: Compute the probabilities for each value of Y in Example 2.24.

2.11 Numerical Events & RV's

Example 2.24 Solution:

Let H and T represent head and tail, respectively; and let an ordered pair of symbols identify the outcome for the first and second coins. (Thus, HT implies a head on the first coin and a tail on the second.) Then the four sample points in S are E_1: HH, E_2: HT, E_3: TH and E_4: TT. The values of Y assigned to the sample points depend on the number of heads associated with each point. For E_1: HH, two heads were observed, and E_1 is assigned the value $Y = 2$. Similarly, we assign the values $Y = 1$ to E_2 and E_3 and $Y = 0$ to E_4. Summarizing, the random variable Y can take three values, $Y = 0, 1$, and 2, which are events defined by specific collections of sample points:

$$\{Y = 0\} = \{E_4\}, \qquad \{Y = 1\} = \{E_2, E_3\}, \qquad \{Y = 2\} = \{E_1\}.$$ ■

Example 2.25 Solution:

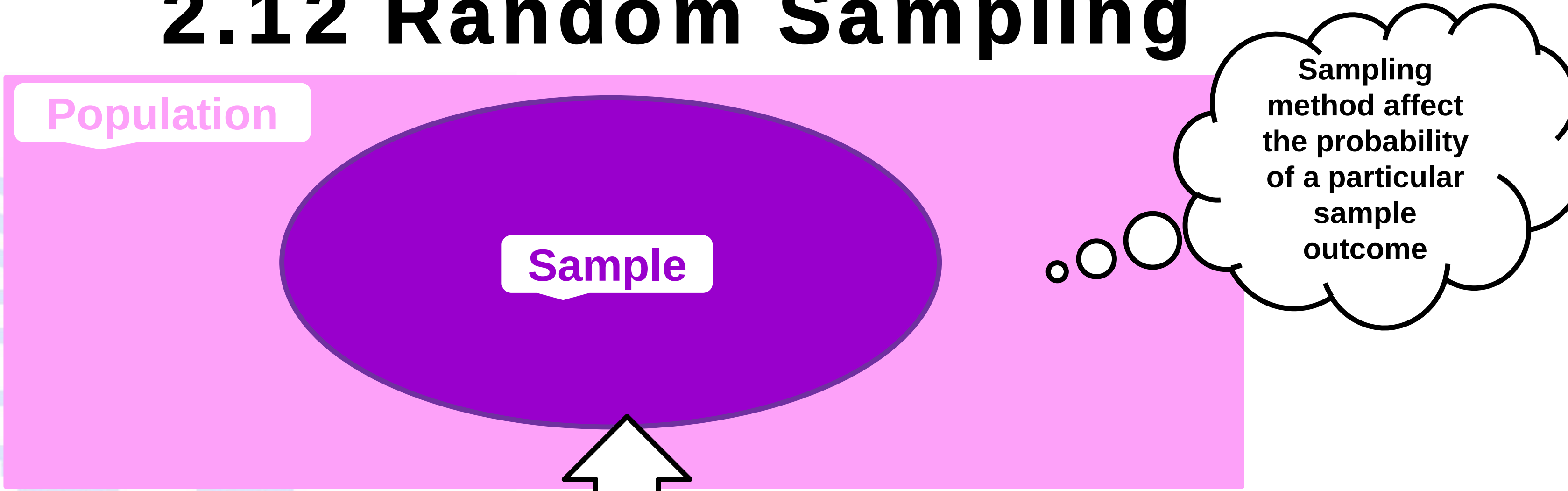
The event $\{Y = 0\}$ results only from sample point E_4. If the coins are balanced, the sample points are equally likely; hence,

$$P(Y = 0) = P(E_4) = 1/4.$$

Similarly,

$$P(Y = 1) = P(E_2) + P(E_3) = 1/2 \quad \text{and} \quad P(Y = 2) = P(E_1) = 1/4.$$ ■

2.12 Random Sampling



Statistical Experiment:
Select sample from population
Sampling without & with replacement

Can the number elements ⊂ sample > number of elements ⊂ population?

Definition 2.13:

Let N and n represent the numbers of elements in the population and sample, respectively. If the sampling is conducted in such a way that each of the $\binom{N}{n}$ samples has an equal probability of being selected, the sampling is said to be random, and the result is said to be a *random sample*.

Method of sampling ~ Design of the Experiment

2.13 Summary

- Concepts (*E*, *Simple E*, *S* & the probability axioms) - provided a probabilistic model(process followed) for calculating the probability of an *E.*
- Inherent in the model is: **Sample-point approach** (Sec.2.5) & **Counting rules** useful in applying the sample-point method (Sec. 2.6).
- Concepts (conditional probability, operations of set algebra, laws of probability): **Event-composition method** (Sec 2.9).
- **Law of Total Probability and Bays Rule**

Theory of probability provides the theory & tools for calculating the **probabilities of numerical events &** hence the **probability distributions** for the **random variables**

STA211 Hand Written Theory Report 1: Chapter 2 and 3

Due Date: Latest Thursday, 5 March 2020
Due Time: At the end of your tutorial – Period 5 at 12h55.

NB

- Please make sure your **student number** is written on your submission paper.
- You are required to write out your theory report by hand in ink pen.
- Make sure your handwriting is legible.
- Make sure that you submit your own written work as found on the presentations or in the in the prescribed textbook. (read section 3.5 in the General Calendar Part 1. It deals with matters of plagiarism and academic dishonesty)
- **No** late submissions will be accepted.
- The submission of **both** **Theory Report 1 and Theory Report 2** can be used to replace the two weakest or two "missed/absent" tutorial test marks obtained.
- Please ensure that you **sign** the theory report **submission register** when submitting your reports.

Instruction

In your own hand writing, write down the:

i) 13 definitions and 9 theorems with proofs (where applicable) found in Chapter 2.

ii) 16 definitions and 14 theorems with proofs (where applicable) found in Chapter 3.

as covered in class lecture periods (see PowerPoint presentations) from the prescribed textbook (Mathematical Statistics with Applications by Wackerly, Mendenhall and Scheaffer 7th Edition).

STA211 Hand Written Theory Report 2: Chapter 4 and 5

Due Date: Thursday, 9 April 2020
Due Time: At the end of your tutorial – Period 5 @ 12h55

NB

• Please make sure your **student number** is written on your submission paper.
• You are required to write out your theory report by hand in ink pen.
• Make sure your handwriting is legible.
• Make sure that you submit your own written work (read section 3.5 in the General Calendar Part 1. It deals with matters of plagiarism and academic dishonesty)
• **No** late submissions will be accepted.
• The submission of **both** **Theory Report 1 and Theory Report 2** can be used to replace the two weakest or two "missed/absent" tutorial test marks obtained.
• Please ensure that you **sign** the theory report **submission register** when submitting your reports

Instruction

In your own hand writing, write down the:

i) 14 definitions and 13 theorems with proofs (where applicable) found in Chapter 4.
ii) 13 definitions and 15 theorems with proofs (where applicable) found in Chapter 5.

as covered in class lecture periods (see PowerPoint presentations) from the prescribed textbook (Mathematical Statistics with Applications by Wackerly, Mendenhall and Scheaffer 7th Edition).