Probability

The aim of this chapter is to revise the basic rules of probability. By the endof this chapter, you should be comfortable with:•conditional probability, and what you can and can't do with conditional expressions;•the Partition Theorem and Bayes' Theorem;•
First-Step Analysis for finding the probability that a process reaches some state, by conditioning on the outcome of the first step;•calculating probabilities for continuous and discrete random variables.
2.1Sample spaces and events Definition: A sample space, Ω, is a set of possible outcomes of a random experiment. Definition: An event, A, is a subset of the sample space. This means that event A is simply a collection of outcomes. Example: Random experiment: Pick a person in this class at random. Sample space: Ω ={all people in class}Event A: A={all males in class}.Definition: Event A occurs if the outcome of the random experiment is a member of the set A. In the example above, event A occurs if the person we pick is male.
172.2Probability Reference List The following properties hold for all events A,B.•P(∅) = 0.•0≤P(A)≤1.•Complement:P(A) = 1−P(A).•Probability of a union: P(A∪B) =P(A)

$+P(B)-P(A\cap B)$.For three events A,B,C:$P(A\cup B\cup C) =P(A)+P(B)+P(C)-P(A\cap B)-P(A\cap C)-P(B\cap C)+P(A\cap B\cap C)$.If A and B are mutually exclusive, then $P(A\cup B) =P(A) +P(B)$.•Conditional probability: $P(A|B) =P(A\cap B)P(B)$.•Multiplication rule: $P(A\cap B) =P(A|B)P(B) =P(B|A)P(A)$.•The Partition Theorem:if$B_1, B_2, . . . , B_m$ form a partition of Ω, then $P(A) =_m\sum_{i=1}P(A\cap B_i) =_m\sum_{i=1}P(A|B_i)P(B_i)$ for any event A. As a special case, B and B partition Ω, so: $P(A) =P(A\cap B) +P(A\cap B)=P(A|B)P(B) +P(A|B)P(B)$ for any A,B.• Bayes' Theorem: $P(B|A) =P(A|B)P(B)P(A)$.More generally, if$B_1, B_2, . . . , B_m$ form a partition of Ω, then $P(B_j|A) =P(A|B_j)P(B_j)\sum_{mi=1}P(A|B_i)P(B_i)$for any j.• Chains of events: for any events$A_1, A_2, . . . , A_n$,$P(A_1\cap A_2\cap . . .\cap A_n) =P(A_1)P(A_2|A_1)P(A_3|A_2\cap A_1). . .P(A_n|A_{n-1}\cap . . .\cap A_1)$.

2.3Conditional Probability Suppose we are working with sample space Ω ={people in class}. I want to find the proportion of people in the class who ski. What do I do? Count up the number of people in the class who ski, and divide by the total number of people in the class. P(person skis) =number of skiers in class total number of people in class. Now suppose I want to find the proportion of females in the class who ski. What do I do? Count up the number of females in the class who ski, and divide by the total

number of females in the class. P(female skis) =number of female skiers in class total number of females in class. By changing from asking about everyone to asking about females only, we have:•restricted attention to the set of females only, or: reduced the sample space from the set of everyone to the set of females, or: conditioned on the event{females}.We could write the above as: P(skis| female) =number of female skiers in class total number of females in class. Conditioning is like changing the sample space: we are now working in a new sample space of females in class.
19 In the above example, we could replace ‘skiing’ with any attribute B. We have: P(skis) =# skiers in class# class; P(skis| female) =# female skiers in class# females in class; so: P(B) =#B’s in class total # people in class, and: P(B| female) =# female B’s in class total # females in class=# in class who are B and female# in class who are female. Likewise, we could replace ‘female’ with any attribute A:P(B|A) =number in class who are B and A number in class who are A. This is how we get the definition of conditional probability: P(B|A) =P(B and A)P(A).=$P(B \cap A)P(A)$.By conditioning on event A, we have changed the sample space to the set of A’s only. Definition: Let A and B be events on the same sample space: so $A \subseteq \Omega$ and $B \subseteq \Omega$. The conditional

probability of event B, given event A, is $P(B|A) = P(B \cap A) P(A)$.

20 Multiplication Rule:(Immediate from above). For any events A and B, $P(A \cap B) = P(A|B)P(B) = P(B|A)P(A) = P(B \cap A)$. Conditioning as ‘changing the sample space’ The idea that“ conditioning” = “changing the sample space” can be very helpful in understanding how to manipulate conditional probabilities. Any ‘unconditional’ probability can be written as a conditional probability: $P(B) = P(B|\Omega)$. Writing $P(B) = P(B|\Omega)$ just means that we are looking for the probability of event B, out of all possible outcomes in the set Ω. In fact, the symbol P belongs to the set Ω: it has no meaning without Ω. To remind ourselves of this, we can write $P = P_\Omega$. Then $P(B) = P(B|\Omega) = P_\Omega(B)$.Similarly, $P(B|A)$ means that we are looking for the probability of event B, out of all possible outcomes in the set A. So A is just another sample space. Thus we can manipulate conditional probabilities $P(\cdot|A)$just like any other probabilities, as long as we always stay inside the same sample space A. The trick: Because we can think of A as just another sample space, let’s write $P(\cdot|A) = P_A(\cdot)$Note: NOT standard notation! Then we can use P_A just like P, as long as we remember to keep the A subscript on EVERY P that we write.

21 This helps us to make quite complex manipulations of conditional probabilities without

thinking too hard or making mistakes. There is only one rule you need to learn to use this tool effectively: $P_A(B|C) = P(B|C \cap A)$ for any A,B,C.(Proof: Exercise).The rules: $P(\cdot|A) = P_A(\cdot)$ $P_A(B|C) = P(B|C \cap A)$ for any A,B,C.