6.5 Combining Probabilities (part 2)

- Independent Events vs. Dependent Events
  - Is your probability of getting an "A+" grade related to studying? Or are those two events unrelated?
- "And" probabilities
  - We call the probability of event A *and* event B occurring a *joint probability*.
- What is the probability that event A occurs *given* that I know event B has occurred?
  - This relates to a *conditional probability*.

Independent Events

- Independent events – Two events are independent if the outcome of one event does not influence the probability of the other event.
  - A) Coin flips as Head on first toss
  - B) Coin flips as Head on second toss
    - → INDEPENDENT: The outcome of a second coin flip does not depend on the outcome of the first coin flip.
  - A) You have red hair **and** B) You get an A+.
    - → INDEPENDENT: Your hair color is not related to how well you do in the class.

Dependent Events

- Dependent events - Two events are dependent if the outcome of one event influences or affects the probability of the other event.
  - A) You turn-in homework
  - B) You get an A+.
    - → NOT INDEPENDENT: The chance that you get an A+ increases if you turn-in your homework.
  - A) It snows **and** B) Class is cancelled
    - → NOT INDEPENDENT: The chance of B) depends on whether A) occurred.

Independent Events

- The probability of the joint occurrence of independent events involves the **multiplication** of each respective probability.
- Example:

  A) Rolling a 4 **and** then B) Rolling a 6

  - These two events are independent.
  - Rolling a 4 on the first roll does not affect the chance of rolling a 6 on the second roll.
  - P(A)=1/6 and P(B)=1/6
  - Because these events are independent…

P(A and B) = P(A) x P(B)

Independent Events

- Example: A) Rolling a 4 **and** then B) Rolling a 6
  - P(A and B) = P(A) x P(B)
  - P(rolling a 4 on the 1st and a 6 on the 2nd roll)

    = P(rolling a 4 on the 1st) x P(rolling a 6 on the 2nd)

    =1/6 x 1/6=1/36

NOTE: Using our earlier visual method of rolling two dice, we know there are 36 equally likely outcomes from rolling two dice, and only (4,6) qualifies under the above. Thus, we can verify a 1/36 chance of this sequence of events.

Probability Rules – Independent Events

- Multiplication rule for independent events.

  P(1st outcome **and** 2nd outcome) =

  P(1st outcome) x P(2nd outcome)

Example

P(flip a 'head' **and** flip a 'head')= $\frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$

NOTE: By our earlier method, we know there are 4 equally likely outcomes from tossing two coins (HH,TH,HT,TT), and only HH qualifies under the above. Thus, we can verify a 1/4 chance of this sequence of events.

“*And*” Probabilities

Independent Events

“*And*” Probability for Independent Events

Two events are **independent** if the outcome of one event does not affect the probability of the other event. Consider two independent events A and B with probabilities $P(A)$ and $P(B)$. The probability that A and B occur together is

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

This principle can be extended to any number of independent events. For example, the probability of A, B, and a third independent event C is

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

Independent Events

- Example:
  - A bag contains 4 blue and 5 red chips.
  - A coin is in your pocket.
  - A deck of cards is on the table.
- Suppose you randomly select a chip, then flip the coin, then randomly select a card.
- Find the probability of getting…
  - A blue chip AND a Head AND a ♣.
  - ANS: 4/9 x 1/2 x 1/4 = 1/18 = 0.0556

Dependent Events

- If the events are dependent, we can not just multiply their 'unconditional' probabilities to get their *joint probability* (which is the probability that they both occur).

- To compute the *joint probability*, we will use a conditional probability. The following notations all represent the same **conditional probability:**
  - P(event A occurs given B occurs)
  - P(A given B)
  - $P(A|B)$

TIME OUT TO THINK

Without doing calculations, compare the probability of drawing three hearts in a row when the card is replaced and the deck is shuffled after each draw to the probability of drawing three hearts in a row without replacement. Which probability is larger and why?

Example: Choosing without replacement

- A bag contains: 5 red, 5 green, 5 blue chips
- Suppose we draw two chips **WITHOUT** replacement.
  - What is the probability of drawing a blue on the first draw?
  - What is the probability of drawing a blue on the second draw?

Example: Choosing without replacement

- A bag contains: 5 red, 5 green, 5 blue chips
- Suppose we draw two chips **WITHOUT** replacement.
  - What is the probability of drawing a blue on the first draw?

    5/15=1/3=0.3333

  - What is the probability of drawing a blue on the second draw?

    we might say “it depends”…

Example: Choosing without replacement

- The probability of drawing a blue on the second draw DEPENDS on what you got on the first draw.
  - P(blue on 2^{nd}|blue on 1^{st})=4/14
  - P(blue on 2^{nd}|not blue on 1^{st})=5/14

Either way, there's only 14 left in the bag for the second draw.

Because the outcome of the first event affects the probability of the second event, then we say these events (the first and second draws) are **dependent events**.

Dependent Events

- Computing *the joint probability* of dependent events still involves multiplication, but we have to consider how prior events affect subsequent events.

Example (continued):

- A bag contains: 5 red, 5 green, 5 blue chips
- What is the probability of drawing two blue chips when drawing **without** replacement?
  - P(blue on 1st **and** blue on 2nd)

    "Given"

    = P(blue on 1st) x P(blue on 2nd|blue on 1st)

    = 5/15 x 4/14 = 20/210 = 0.0952

- For <u>dependent</u> events A and B,

Conditioning notation

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

Example (continued):

- For <u>dependent</u> events A and B,

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

or

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

"*And*" Probability for Dependent Events

Two events are **dependent** if the outcome of one event affects the probability of the other event. The probability that dependent events *A* and *B* occur together is

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

where *P*(*B* given *A*) means the probability of event *B* given the occurrence of event *A*.

This principle can be extended to any number of individual events. For example, the probability of dependent events *A*, *B*, and *C* is

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

"And" Probability for Dependent Events

When we have more the two dependent events, we can 'group' some of the events to use our earlier definition from the two events scenario to get…

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

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

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

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

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

This is a very useful technique or calculating complicated probabilities.

Example: 3 draws **without replacement** from a deck of cards

- What is the probability of getting 3 jacks when drawing 3 cards?
  - $P(J_1 \text{ and } J_2 \text{ and } J_3)$

$$= P(J_1) \times P(J_2|J_1) \times P(J_3|J_1 \text{ and } J_2)$$

$$= 4/52 \times 3/51 \times 2/50$$

$$= 0.000181$$

Depends on what happened on previous draws

Back to *independent* events…

Example: 3 draws **with replacement** from a deck of cards

- What is the probability of getting 3 jacks when drawing 3 cards and replacing the card before each subsequent draw?
  - $P(J_1 \text{ and } J_2 \text{ and } J_3) = P(J_1) \times P(J_2) \times P(J_3)$

    $= 4/52 \times 4/52 \times 4/52$

    $= 0.000455$

Does NOT depend on what happened on previous draws

Can we *ever* treat dependent events as independent events?

- Consider the poker chip example, but let's increase the number of chips:
  - 5000 red, 5000 green, 5000 blue chips

- What is the probability of drawing two blue chips when drawing **without** replacement?
  - P(blue on 1^{st} **and** blue on 2^{nd})

    = P(blue on 1^{st}) x P(blue on 2^{nd}|blue on 1^{st})

    = 5000/15,000 x 4999/14,999

    = 0.111096

Can we *ever* treat dependent events as independent events?

- If we considered these events to be independent, we would compute…
  - P(blue on 1^{st} **and** blue on 2^{nd})

    = 5000/15,000 x 5,000/15,000

    = 0.111111

  which is nearly identical to the computation based on dependent events.

Can we *ever* treat dependent events as independent events?

- In general, when you draw only a few items from a VERY LARGE population <u>without replacement</u>, the computed probabilities are very similar to those based on the <u>with replacement</u> framework.

- In this case, we can essentially treat dependent events as independent.
  - The book gives a guideline of sampling no more than 5% (small portion) of the overall population.

Can we *ever* treat dependent events as independent events?

- Jury selection:
  - We draw a juror (**without** replacement) from a very large population with equal number of men and women.

    Technically, subsequent draws are not independent, but we can essentially treat them as independent.

  - Find the probability of selecting 9 male jurors.

    $P(9 \text{ males}) = 0.5 \times 0.5 \times \ldots \times 0.5 = (0.5)^9 = 0.00195$

Example: Salary and Color of car

- Do people with high salaries drive red cars?
- A company has 350 employees. The frequency table for Salary (high or low) and Color of car (red or not red) are shown in the table below:

Car color

		red	**not red**	total
Salary	**low**	28	252	280
	high	7	63	70
	total	35	315	350

Example: Salary and Color of car

- Suppose we randomly choose an employee.
  - **Let H be the event of high salary.**
    - P(H)=70/350 = 0.20
  - **Let R be the event of red car.**
    - P(R)=35/350=0.10

Are these independent events?

Car color

		red	**not red**	total
	low	28	252	280
Salary	**high**	7	63	70
	total	35	315	350

Example: Salary and Color of car

- The probability that a randomly drawn employee is in the high salary group is 0.20
- What if I randomly choose from the red car group? Do I have a higher chance of getting a high salary employee?
  - P(H|R)=7/35=0.20
  - It's the same!!!

		Car color		
		red	**not red**	total
	low	28	252	280
Salary	**high**	7	63	70
	total	35	315	350

Example: Salary and Color of car

- So, $P(H|R)=0.20$ and $P(H)=0.20$

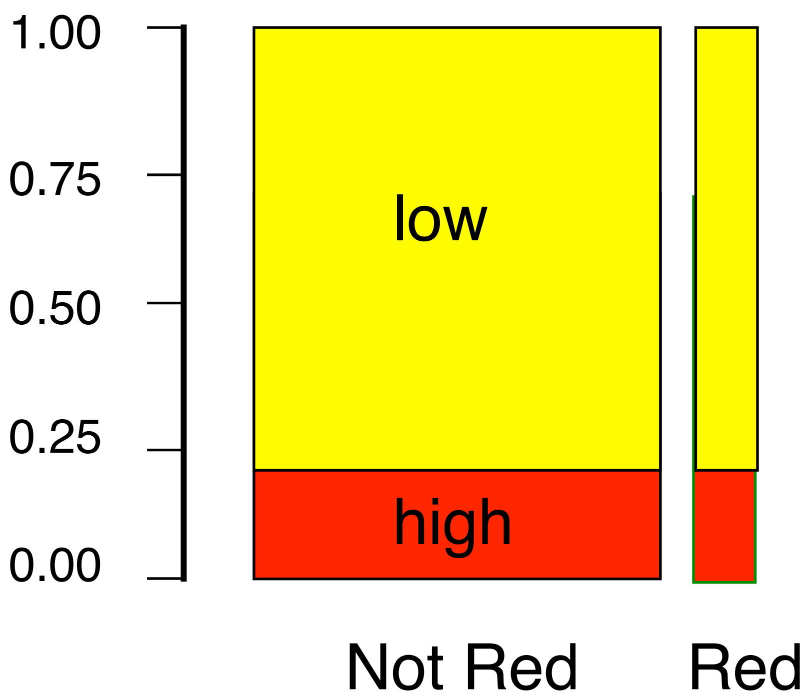
$$P(H|R)=P(H)=0.20$$

- These two events are independent.

- If I'm picking an employee at random, knowing that they drive a red car doesn't help me in predicting if they will have a high salary. It's a 20% chance regardless of the color of car they drive (red car or not).

Independent Events

(mosaic plot)



$P(R)=0.10$

$P(\overline{R})=0.90$

Complement of R, which means 'not R'.

$P(H)=0.20$

$P(H|R)=0.20$

This says they're independent.

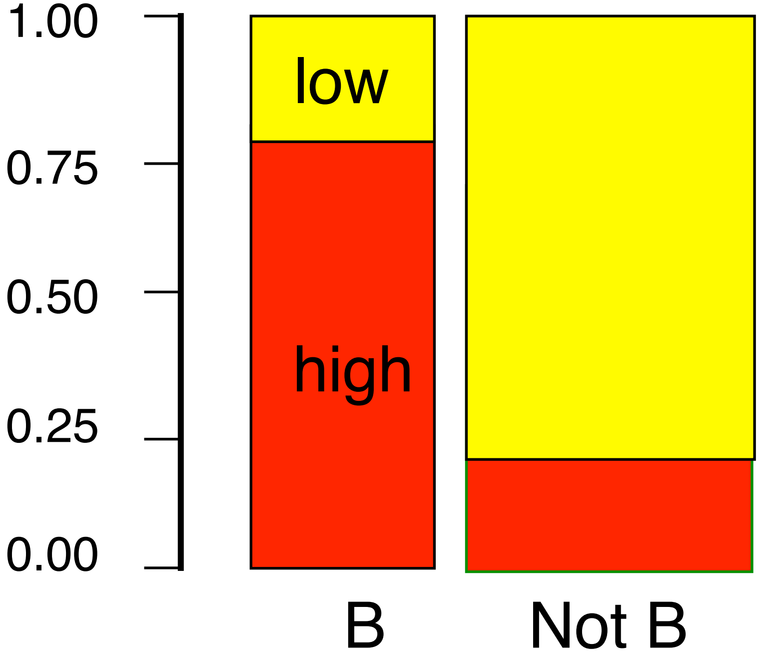
Regardless of whether a person has a red car (R) or not ($\overline{R}$), there's still a 0.20 probability that they have a high salary.

Independent Events

- Two events are independent if any one of the following equivalent statements is true:
  - $P(B|A)=P(B)$
  - $P(A|B)=P(A)$
  - $P(A \text{ and } B)=P(A) \times P(B)$

Dependent Events

B: Person has B.S. degree
H: Person has a high salary



$P(B)=0.30$

$P(\overline{B})=0.70$

$P(H|B)=0.80$

$P(H)=0.40$

This says they're NOT independent. $P(H|B)\neq P(H)$

The probability of having a high salary **depends on** whether a person has a B.S. degree (B) or not ($\overline{B}$).

Summary

Table 6.11 provides a summary of the formulas we've used in combining probabilities.

Table 6.11 Summary of Combining Probabilities

And **probability: independent events**	***And*** **probability: dependent events**	***Either/or*** **probability: non-overlapping events**	***Either/or*** **probability: overlapping events**
$P(A \text{ and } B) = P(A) \times P(B)$	$P(A \text{ and } B) = P(A) \times P(B \text{ given } A)$	$P(A \text{ or } B) = P(A) + P(B)$	$P(A \text{ or } B) = P(A) + P(B) - P(A \text{ and } B)$