Probability Rules!

November 6, 2013

Jason Bryer
epsy530.bryer.org

Or but NOT disjoint

Your Wallet

- S = {1,2, 5,10, 20,50, $100}
- A = {odd numbered value} = {1,5}
- B = {bill with a building} = {5,10, 20,50, $100}

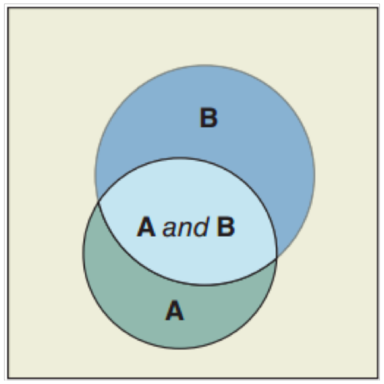
Why is P(A or B) ≠ P(A) + P(B)?

- Answer: A and B are not disjoint.
- The intersection A and B = {$5} is double counted.
- To find P(A or B), subtract P(A and B).

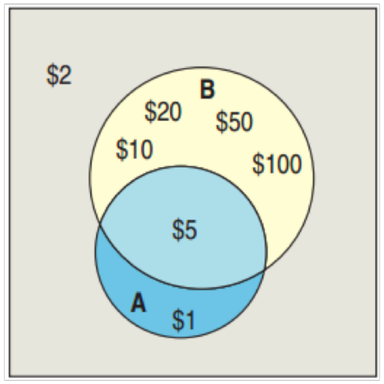
The General Addition Rule

$$P(A \quad or \quad B) = P(A) + P(B) – P(A \quad and \quad B)$$

Add the probabilitiesof the two events and then subtract the probability of their intersection.



P(odd amount or bill with a building) =
P(A) + P(B) – P(A and B} =
P({1, 5}) + P({5, 10, 20, 50, 100}) – P({5})



General Addition Rule Example

Survey

- Are you currently in a relationship?
- Are you involved in sports?

Results

- 33% are in a relationship.
- 25% are involved in sports.
- 11% answered “yes” to both.

Problem

- Find the probability that a randomly selected student is in a relationship or is involved in sports.

General Addition Rule Example

Survey

- Are you currently in a relationship?
- Are you involved in sports?

Results

- 33% are in a relationship.
- 25% are involved in sports.
- 11% answered “yes” to both.

Problem

- Find the probability that a randomly selected student is in a relationship or is involved in sports.

33% Relationship, 25% Sport, 11% Both

Events

- R = {in a relationship}
- S = {involved in sports} Calculations
- P(R or S) = P(R) + P(S) – P(R and S) = 0.33 + 0.25 – 0.11 = 0.47

Conclusion

- There is a 47% chance that a randomly selected student is in a relationship or is involved sports.

Using Venn Diagrams

P(not in relationship and no sports)

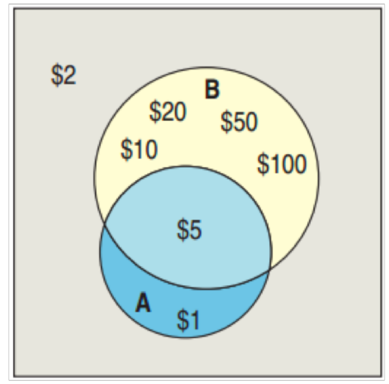
- $P(R^C \quad and \quad S^C)$
- This is the part outside of both circles: 0.53.

P(in a relationship but no sports)

- $P(R \quad and \quad S^C)$
- This is the part in the circle R that is outside S: 0.22.

P(in a relationship or involved in sports but not both)

- $P((R \quad and \quad S^C)$ or $(R^C \quad and \quad S))$
- This is the combination of the circles minus the intersection: 0.22 + 0.14 = 0.36



Contingency Table

A table that displays the results of two categorical questions is called a contingency table.

- P(girl) = 251/478 = 0.525
- P(girl and popular) = 91/478 = 0.190
- P(sports) = 90/478 = 0.188

		Goals			
		Grades	Popular	Sports	Total
Sex	Boy	117	50	60	227
	Girl	130	91	30	251
	Total	247	141	90	478

Conditional Probability

What if we knew the chosen personwas a girl? Wouldthat change the probability that the girl's goal was sports?

- Yes! We write P(sports I girl)
- Only look at Girl row: P(sports I girl) = 30/251 = 0.120
- Find the probability of selecting a boy given the goal is grades.
- P(boy I grades) = 117/247 = 0.474

		Goals			
		Grades	Popular	Sports	Total
Sex	Boy	117	50	60	227
	Girl	130	91	30	251
	Total	247	141	90	478

Conditional Probability Formula

Probability of B given A:

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

		Goals			
		Grades	Popular	Sports	Total
Sex	Boy	117	50	60	227
	Girl	130	91	30	251
	Total	247	141	90	478

$$P(girl \quad | \quad popular) = \frac{P(girl \quad and \quad popular)}{P(popular)}$$

$$P(girl \quad | \quad popular) = \frac{91/478}{141/478}$$

$$P(girl \quad | \quad popular) = \frac{91}{P(141)} = 0.65$$

The General Multiplication Rule

- For A and B independent, we had:
  $P(A \text{ and } B) = P(A) \times P(B)$
- Rearranging the conditional probability equation, we get the General Multiplication Rule:
  $P(A \text{ and } B) = P(A) \times P(B \mid A)$
- Equivalently,
  $P(A \text{ and } B) = P(B) \times P(A \mid B)$

Independence

Events A and B are independent if knowing A happened does not change the probability of B. Symbolically:

- A and B are independent $\leftrightarrow$ $P(B \mid A) = P(B)$

Equivalent formulas for independence:

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

Grades and Girl Independent?

Determine if the goal of good grades and sex are independent.

- P(grades I girl) = 130/251 ≈ 0.52
- P(grades) = 247/478 ≈ 0.52
- To two decimal places, they are independent.

Are the goal of sports and sex independent?

- P(sports I boy) = 60/227 ≈ 0.26
- P(sports) = 90/478 ≈ 0.19
- No, the goal of sports and sex are dependent.

		Goals			
		Grades	Popular	Sports	Total
Sex	Boy	117	50	60	227
	Girl	130	91	30	251
	Total	247	141	90	478

Relationships, Sports, and Independence

33% in a relationship, 25% involved in sports, 11% both

Are being in a relationship and being involved in sports independent?

- P(relationship) = 0.33
- P(sports) = 0.25
- P(relationship and sports) = 0.11
- $0.33 \times 0.25 = 0.0825 \neq 0.11$
- No, they are dependent.

Are they disjoint?

- P(relationship and sports) = $0.11 \neq 0$
- No, they are not disjoint.

Independent ≠ Disjoint

Disjoint events cannot be independent.

Consider the events:

- Course grade A
- Course grade B
- Disjoint: You can't get both.
- Not independent: $P(A \mid B) = 0 \neq P(A)$
- A and B are disjoint (also called mutually exclusive) but not independent.

Marginal and Joint Probabilities

- 73% use e-mail, 62% text, 49% both
- Draw a partial table
- 0.73 and 0.62 are called marginal probabilities.
- 0.49 is a joint probability.

How can we complete the table?

Use Text Messaging	Use E-Mail: Yes	Use E-Mail: No	Total
Yes	0.49		0.62
No			
Total	0.73		1.00

Marginal and Joint Probabilities

- 73% use e-mail, 62% text, 49% both
- Draw a partial table
- 0.73 and 0.62 are called marginal probabilities.
- 0.49 is a joint probability.

How can we complete the table?

- The sum must add up
- 0.49 + ? = 0.73
- 0.62 + ? = 1.00

		Use E-Mail		
		Yes	No	Total
Use Text Messaging	Yes	0.49		0.62
	No			
	Total	0.73		1.00

		Use E-Mail		
		Yes	No	Total
Use Text Messaging	Yes	0.49	0.13	0.62
	No	0.24	0.14	0.38
	Total	0.73	0.27	1.00

73% Use E-mail, 62% Text, 49% Both

1. E-mail and text mutually exclusive?
2. E-mail and text independent?

Plan:

- A = {uses e-mail}
- B = {texts}
- $P(A) = 0.73$
- $P(B) = 0.62$
- $P(A \text{ and } B) = 0.49$

73% Use E-mail, 62% Text, 49% Both

1. E-mail and text mutually exclusive?
2. E-mail and text independent?

Plan:

- A = {uses e-mail}
- B = {texts}
- P(A) = 0.73
- P(B) = 0.62
- P(A and B) = 0.49

1. E-mail and text mutually exclusive?

- P(A and B) = 0.49 ≠ 0
- Conclusion: E-mail and text are not mutually exclusive.

73% Use E-mail, 62% Text, 49% Both

1. E-mail and text mutually exclusive?
2. E-mail and text independent?

Plan:

- A = {uses e-mail}
- B = {texts}
- P(A) = 0.73
- P(B) = 0.62
- P(A and B) = 0.49

1. E-mail and text independent?

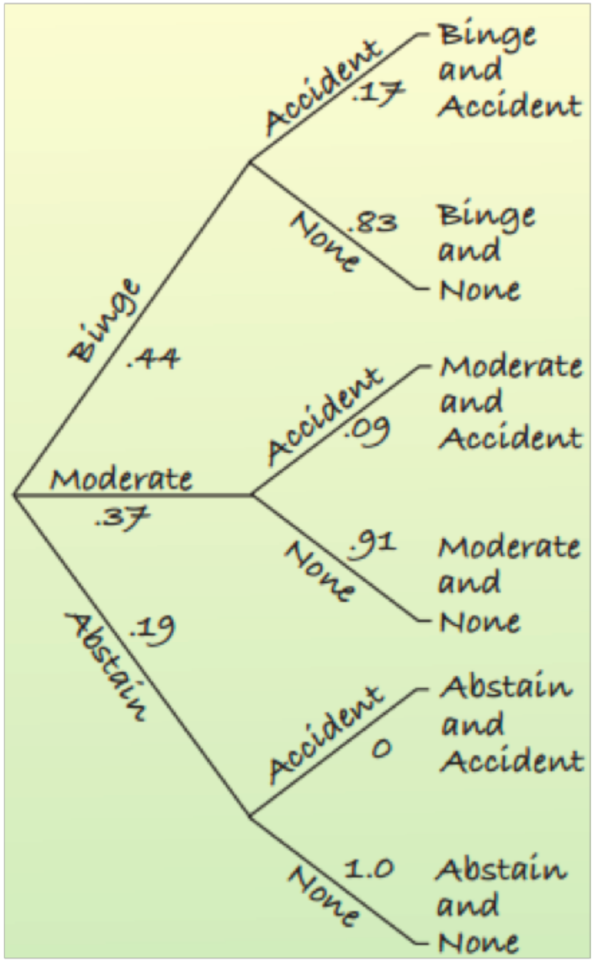
- Make a table

		Use E-Mail		
		Yes	No	Total
Use Text Messaging	Yes	0.49	0.13	0.62
	No	0.24	0.14	0.38
	Total	0.73	0.27	1.00

- P(B I A) = 0.49/0.73 ≈ 0.67
- P(B) = 0.62 ≠ 0.67
- Not independent
- Conclusion: Since the respondents who use e-mail are more likely to text, they are not independent.

Tree Diagrams

- 44% binge drink, 37% drink moderately, 19% don't drink
- Binge drinkers: 17% in an alcohol related accident
- Non-bingers: 9% in an alcohol-related accident
- Find the probability of being a binge drinker and has had an alcohol-related accident.
- Venn diagrams and tables are not great for conditional probabilities.
- Use a tree diagram.



Tree Diagrams

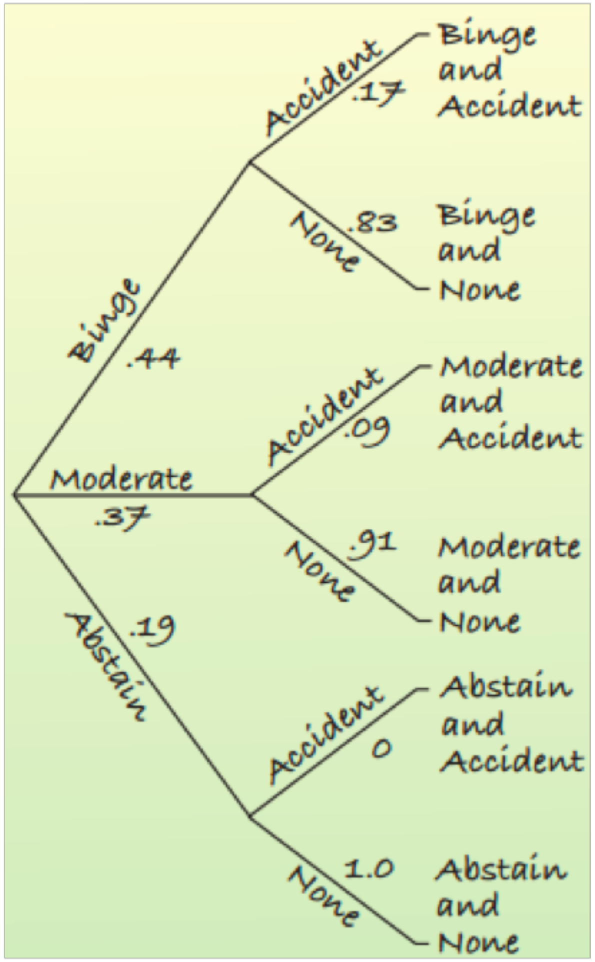
This tree diagram gives the complete information.

Notice the sums:

- 0.17 + 0.83 = 1
- 0.09 + 0.91 = 1
- 0 + 1.0 = 1
- 0.44 + 0.37 + 0.19 = 1

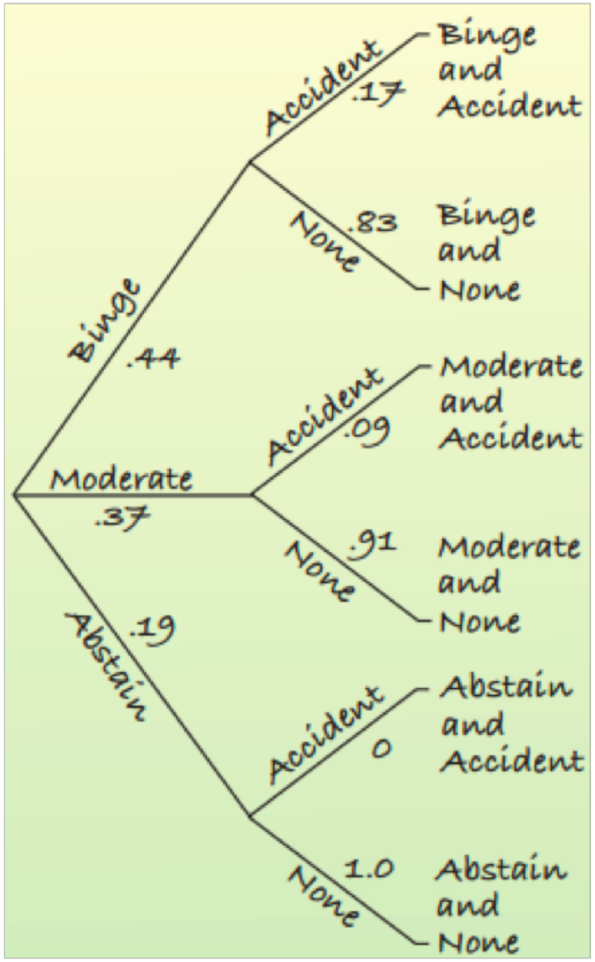
Conditional Probabilities:

- P(none I binge) = 0.83



Probabilities From Trees

- P(moderate and accident) = 0.37 × 0.09 = 0.0333
- P(abstain and accident) = 0.19 × 0 = 0
- P(none) = (0.44 × 0.83) + (0.37 × 0.91) + (0.19 × 1.0) = 0.8919



Tree Diagram Facts

- The sum of the probabilities emanating from any branch is 1.
- The final outcomes are disjoint.
- To find a conditional probability, multiply across.