

Two-Way Tables

Gaston Sanchez

Creative Commons Attribution Share-Alike 4.0 International CC BY-SA

Two-way tables,
crosstables,
contingency tables

Example

Body Image

Random sample of 1200 US college students

“What is your perception of your own body?”

- About right
- Underweight
- Overweight

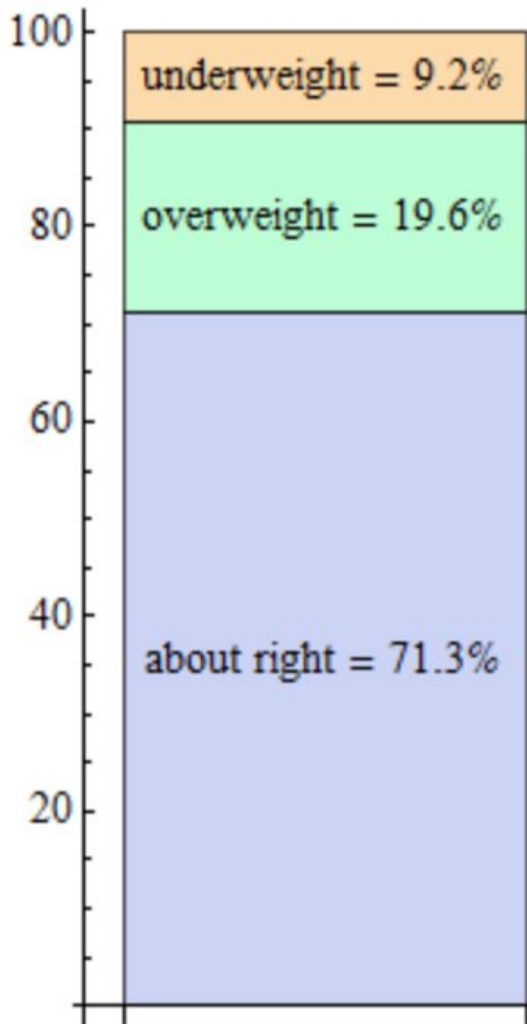
For instance ...

Student	Body Image
...	...
25	overweight
26	about right
27	underweight
28	about right
29	about right
...	...

Frequency Table

Category	Count	Proportion	Percentage
underweight	110	$110/1200 = 0.092$	9.2%
overweight	235	$235/1200 = 0.196$	19.6%
about right	855	$855/1200 = 0.713$	71.3%

Body Image



Gender and Body Image

Student	Gender	Body Image
...
25	M	overweight
26	M	about right
27	F	underweight
28	F	about right
29	M	about right
...

Two-Way Table

	About Right	Overweight	Underweight
Female	560	163	37
Male	295	72	73

Two-Way Table

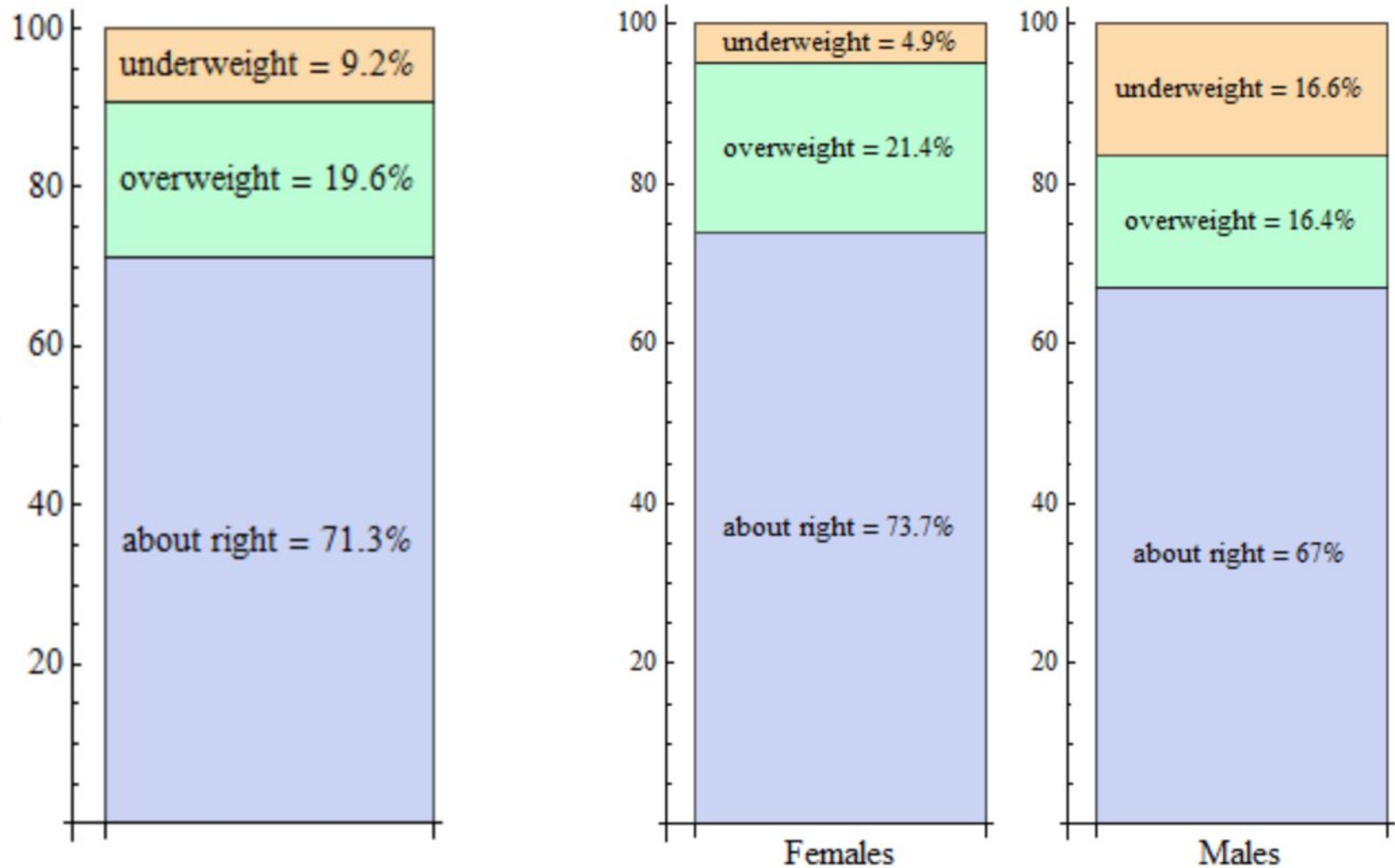
	About Right	Overweight	Underweight	<i>Row Totals</i>
Female	560	163	37	760
Male	295	72	73	440
<i>Column Totals</i>	855	235	110	1200

Two-Way Table

	About Right	Overweight	Underweight	<i>Row Totals</i>
Female	560	163	37	760
Male	295	72	73	440
<i>Column Totals</i>	855	235	110	1200

Is Body Image related to Gender?
(discussed in class)

Is Body Image related to Gender?



Example

Example

Suppose we observe 2 qualitative binary variables:

Gender

male, female

Condition

smoker, non-smoker

Example of crosstable 2x2

Table formed by **crossing** Gender and Condition

	smoker	non-smoker
male	20	35
female	15	40

Example of crosstable 2x2

Table formed by **crossing** Gender and Condition

		B	B ^c
		smoker	non-smoker
A	male	20	35
A ^c	female	15	40

note that these are absolute frequencies

Example of crosstable 2x2

Table formed by **crossing** Gender and Condition

		B	B ^c	
		smoker	non-smoker	Total
A	male	20	35	55
A ^c	female	15	40	55
Total		35	75	110 grand total

note that these are absolute frequencies

Crosstable 2x2: general case

	B	B ^c	Total
A	A and B	A and B ^c	# A
A ^c	A ^c and B	A ^c and B ^c	# A ^c
Total	# B	# B ^c	# N

In order to get
probabilities...

Probability crosstable 2x2

Table formed by **crossing** Gender and Condition

		B	B ^c	
		smoker	non-smoker	Total
A	male			
A ^c	female			
Total				
				grand total

Probability crosstable 2x2

Table formed by **crossing** Gender and Condition

		B	B ^c	
		smoker	non-smoker	Total
A	male	20/110	35/110	55/110
A ^c	female	15/110	40/110	55/110
Total		35/110	75/110	110/110 grand total

note that these are relative frequencies

Probability crosstable 2x2

Table formed by **crossing** Gender and Condition

		B	B ^c	
		smoker	non-smoker	Total
A	male	0.1818	0.3181	0.5
A ^c	female	0.1363	0.3636	0.5
Total		0.3181	0.6818	1.0 grand total

note that these are relative frequencies

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1

$$P(A | B) = P(A \text{ \& } B) / P(B)$$

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1


$$P(A^c | B) = P(A^c \text{ \& } B) / P(B)$$

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1

$$P(A \mid B^c) = P(A \ \& \ B^c) / P(B^c)$$

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1

$$P(A^c \mid B^c) = P(A^c \text{ \& } B^c) / P(B^c)$$

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1

$$P(B \mid A) = P(B \text{ \& } A) / P(A)$$

Probability Crosstable 2x2

	B	B ^c	Total
A	P(A and B)	P(A and B ^c)	P(A)
A ^c	P(A ^c and B)	P(A ^c and B ^c)	P(A ^c)
Total	P(B)	P(B ^c)	1

$$P(B^c \mid A) = P(B^c \& A) / P(A)$$

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1

$$P(B \mid A^c) = P(B \ \& \ A^c) / P(A^c)$$

Probability Crosstable 2x2

	B	B ^c	Total
A	$P(A \text{ and } B)$	$P(A \text{ and } B^c)$	$P(A)$
A ^c	$P(A^c \text{ and } B)$	$P(A^c \text{ and } B^c)$	$P(A^c)$
Total	$P(B)$	$P(B^c)$	1

$$P(B^c \mid A^c) = P(B^c \text{ \& } A^c) / P(A^c)$$

General Crosstable $p \times q$

Crosstables $p \times q$

We observe 2 qualitative variables (*nominal or ordinal*)

X

$A_1, A_2, A_3, \dots, A_p$

Y

$B_1, B_2, B_3, \dots, B_q$

Crosstable $p \times q$: general case

	B_1	B_2	...	B_q	<i>Total</i>
A_1	$A_1 \text{ and } B_1$	$A_1 \text{ and } B_2$...	$A_1 \text{ and } B_q$	$\# A_1$
A_2	$A_2 \text{ and } B_1$	$A_2 \text{ and } B_2$...	$A_2 \text{ and } B_q$	$\# A_2$
A_3	$A_3 \text{ and } B_1$	$A_3 \text{ and } B_2$...	$A_3 \text{ and } B_q$	$\# A_3$
...					...
A_p	$A_p \text{ and } B_1$	$A_p \text{ and } B_2$...	$A_p \text{ and } B_q$	$\# A_p$
<i>Total</i>	$\# B_1$	$\# B_2$...	$\# B_q$	$\# N$

Example

Crosstable for current enrollment in public and private schools by level of education

	Public	Private	<i>Total</i>
Elementary	20%	30%	50%
High School	15%	20%	35%
College	10%	5%	15%
<i>Total</i>	45%	55%	<i>100%</i>

Probability that a student randomly selected is enrolled in Elementary and High School?

$P(\text{enrolled in Elementary and HS}) = ?$

	Public	Private	<i>Total</i>
Elementary	20%	30%	50%
High School	15%	20%	35%
College	10%	5%	15%
<i>Total</i>	45%	55%	100%

Independent?
Mutually Exclusive?
None of the above?

$P(\text{enrolled in Elementary and HS}) = ?$

	Public	Private	<i>Total</i>
Elementary	20%	30%	50%
High School	15%	20%	35%
College	10%	5%	15%
<i>Total</i>	45%	55%	<i>100%</i>

mutually exclusive events

Keep in mind ...

Mut. Exclusive
events

\neq

Independent
events

*typically has to do with
outcomes of same
experiment*

*typically has to do with
outcomes of
different experiments*