Probability Basics
Part 3: Types of Probability

INFO-1301, Quantitative Reasoning 1

University of Colorado Boulder

September 30, 2016

Prof. Michael Paul

Prof. William Aspray

Example

A large government survey of Americans determined that their health status is as follows:

Excellent	Very Good	Good	Fair	Poor
72,867	109,066	88,792	31,568	10,575

Let X be an individual's health status

$P(X = \text{Excellent}) = ?$

$= 72{,}867 / (72{,}867 + 109{,}066 + 88{,}792 + 31{,}568 + 10{,}575)$

$= 0.2329$

Example

Now let's split this into two rows that separate individuals based on whether they have insurance:

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

Let X be an individual's health status

Let Y be an individual's insurance status

$\underbrace{P(X = \text{Excellent} \mid Y = \text{No})} = ?$

The probability that X = Excellent is true, given that Y = No is true.

Example

Now let's split this into two rows that separate individuals based on whether they have insurance:

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

Let X be an individual's health status

Let Y be an individual's insurance status

$P(X = \text{Excellent} \mid Y = \text{No}) = ?$

$= 7{,}196 / (7{,}196 + 11{,}388 + 13{,}359 + 6{,}007 + 1{,}564)$

$= 0.1821$

Different Types of Probability

Let X be the health status of an individual and Y be the insurance status of the individual

P(X = Excellent) **marginal** probability

The probability of exactly one outcome is sometimes called a **marginal** probability

Different Types of Probability

Let X be the health status of an individual and Y be the insurance status of the individual

P(X = Excellent)	**marginal** probability
P(X = Excellent, Y = Yes)	**joint** probability

The probability that two or more outcomes are all true is called a **joint** probability

- This example is equivalent to writing P(X = Excellent AND Y = Yes)

Different Types of Probability

Let X be the health status of an individual and Y be the insurance status of the individual

$P(X = \text{Excellent})$	**marginal** probability
$P(X = \text{Excellent}, Y = \text{Yes})$	**joint** probability
$P(X = \text{Excellent} \mid Y = \text{Yes})$	**conditional** probability

The probability of an outcome, given that one or more other outcomes are true, is a **conditional** probability

- In this example, we would say that the probability of X is *conditioned* on Y

Different Types of Probability

Let X be the health status of an individual and Y be the insurance status of the individual

$P(X = \text{Excellent})$	**marginal** probability
$P(X = \text{Excellent}, Y = \text{Yes})$	**joint** probability
$P(X = \text{Excellent} \mid Y = \text{Yes})$	**conditional** probability

If you know the value of two of these types of probabilities, you can calculate the third

Joint Probability

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

The joint probability of any two outcomes can be calculated from the value in the corresponding cell

$P(X = \text{Excellent},\ Y = \text{No}) = ?$

$= 7{,}196 \,/\, (7{,}196 + 11{,}388 + 13{,}359 + 6{,}007 + 1{,}564 + 65{,}671 + 97{,}708 + 75{,}401 + 25{,}561 + 9{,}042)$

$= 0.0230$

Marginal Probability

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

What if you want the probability of one outcome but you only have a table of joint outcomes?

$P(X = \text{Excellent}) = ?$

$= (7{,}196 + 65{,}671) /$
$(7{,}196 + 11{,}388 + 13{,}359 + 6{,}007 + 1{,}564 + 65{,}671 + 97{,}708 + 75{,}401 + 25{,}561 + 9{,}042)$

$= 0.2329$

Marginal Probability

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

What if you want the probability of one outcome but you only have a table of joint outcomes?

$P(X = \text{Excellent}) = ?$

$= P(X = \text{Excellent},\ Y = \text{No}) + P(X = \text{Excellent},\ Y = \text{Yes})$

Marginalization: The marginal probability of an outcome can be calculated by summing over all joint probabilities that include the outcome

Conditional Probability

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

Conditional probabilities use the proportions within the row or column corresponding to the condition

$P(X = \text{Excellent} \mid Y = \text{No}) = ?$

$= 7{,}196 / (7{,}196 + 11{,}388 + 13{,}359 + 6{,}007 + 1{,}564)$

$= 0.1821$

Conditional Probability

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

Conditional probabilities can also be estimated by dividing the joint probability by the marginal probability of the condition

$P(X = \text{Excellent} \mid Y = \text{No}) = ?$

$= P(X = \text{Excellent}, Y = \text{No}) / P(Y = \text{No})$

$= 0.0230 / 0.1262 = 0.1822$

Summary of Rules

For any two random variables X and Y with values a and b:

$$P(X = a) = \sum_b P(X = a, Y = b)$$

$$P(X = a, Y = b) = P(X = a \mid Y = b) \times P(Y = b)$$

$$P(X = a \mid Y = b) = P(X = a, Y = b) / P(Y = b)$$

Revisiting dice

Suppose you roll two dice. X is the outcome of the first and Y is the outcome of the second.

$P(X = 3) = 1/6$

$P(X = 3 \mid Y = 3) = 1/6$

Knowing the outcome of one die doesn't tell you anything about the other

- $P(X = 3 \mid Y = 3) = P(X = 3)$

Independence

Two random variables are **independent** if knowing the outcome of one does not change the probability of the other

If X and Y are independent then:

$P(X = a \mid Y = b) = P(X = a)$

$P(Y = b \mid X = a) = P(Y = b)$

Independence

Two random variables are **independent** if knowing the outcome of one does not change the probability of the other

If X and Y are independent then:

$$P(X = a, Y = b) = P(X = a) \times P(Y = b)$$

Revisiting dice

Suppose you roll two dice. X is the outcome of the first and Y is the outcome of the second.

$P(X = 3,\ Y = 3)$

$= P(X = 3) \times P(Y = 3)$

$= 1/6 \times 1/6$

$= 1/36$

Independence

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

$P(X = \text{Excellent},\ Y = \text{No}) = 0.0230$

$= 7{,}196 / (7{,}196 + 11{,}388 + 13{,}359 + 6{,}007 + 1{,}564 + 65{,}671 + 97{,}708 + 75{,}401 + 25{,}561 + 9{,}042)$

Independence

	Excellent	Very Good	Good	Fair	Poor
No	7,196	11,388	13,359	6,007	1,564
Yes	65,671	97,708	75,401	25,561	9,042

$\mathrm{P}(X = \text{Excellent},\ Y = \text{No}) = 0.0230$

$= \mathrm{P}(X = \text{Excellent}) \times \mathrm{P}(Y = \text{No})$??

$= 0.2329 \times 0.1262 = 0.0294$

Nope: X and Y are **not** independent