

warm-up exercise: Probability & Notation SOLUTION

This is a very dense summary of basic notions in probability.

Here we use "Pr()" for Probability. In many textbooks and articles you see P() or p().

Examples:

Notation: **Pr(X=5)** = 0.3

Meaning: The **probability** that the random variable X takes the value 5 is 30%.

Notation: **Pr(X)**

Formulas or calculations are often carried out "for any value", i.e. without inserting a concrete value for X. In these cases, one sees the notation Pr(X). Only at the end of a calculation or when applying a formula a concrete value is inserted.

Notation: **E[X]**

Meaning: **Expected Value** (fr: Espérance mathématique, de: Erwartungswert) of a random variable X.

Example: Let X be the number of points of a die. The expected value is $E[X] = 3.5$

Notation: **Pr(X=2, Y=99)** = 0.07

Meaning: The **joint probability** (fr: probabilité conjointe, de: gemeinsame Wahrscheinlichkeit) that the random variable X takes the value 2 and the random variable Y has the value 99 is 7%.

Note that $Pr(X,Y) = Pr(Y,X)$

Notation: $Pr(X) = \sum_Y Pr(X,Y)$ or $Pr(Y) = \sum_X Pr(X,Y)$

Meaning: The **marginal probability** (fr: Probabilité marginale, de: Randverteilung) is calculated by summing the joint distribution over all values of one variable. This calculation eliminates that variables. In the formula above the marginal distribution Pr(X) or Pr(Y) is obtained from the joint probability Pr(X,Y). People use different expressions for marginalization: "Y is summed away", "Y is eliminated", "X is marginalized", etc.

Notation: $Pr(X=3|Y=1)$ = 0.12

Meaning: The **conditional probability** (fr: probabilité conditionnelle, de: bedingte Wahrscheinlichkeit) that the random variable X takes the value 3, if we know (**given**) that the random variable Y takes the value 1, is 12%.

Note that $Pr(X|Y) \neq Pr(Y|X)$!

Joint probability and conditional probability are related:

$$Pr(X,Y) = Pr(X|Y) * Pr(Y)$$

Bayes Rule:

Bayes Rule connects two conditional probabilities:

$$Pr(A|B) = \frac{Pr(B|A)Pr(A)}{Pr(B)}$$

Decision Making under Uncertainty, week 6, warmup

Exercise 1, Repetition Probabilities

This is a "Pen and Paper" exercise. Make sure to write the formulas, not just results.

A medical study shows a correlation between certain blood values and a person's state of health. The following table shows such a (simplified) correlation as "Joint Probability" $\Pr(X,Y)$:

		blood value 0-20	blood value 21-50	blood value 51-100	
		X=1	X=2	X=3	
healthy	Y=1	0.50	0.20	0.05	0.75
ill	Y=2	0.05	0.05	0.15	0.25
		0.55	0.25	0.20	

$\Pr(Y=1)=0.75$
 $\Pr(Y=2)=0.25$

- a) What is the probability that a randomly selected person has a blood value between 51 and 100 and is healthy, i.e. $\Pr(X=3, Y=1)$?

$\Pr(X=3, Y=1) = 0.05.$ joint probability, value from the table

- b) Complete all marginal probabilities directly in the table.

- c) What is the probability that a randomly selected person is ill?

$\Pr(Y=2) = 0.25.$ marginal probability. "X is summed away"

- d) What is the probability that a randomly selected person has a blood value between 21 and 100?

$\Pr(X=2 \text{ or } X=3) = 0.25 + 0.20 = 0.45$

The doctor knows with certainty that a person is ill. Calculate the conditional probability that the blood value of that person is between 0 and 20. Use the table above and the definition $\Pr(X,Y) = \Pr(X|Y) * \Pr(Y)$. Write down the formula, not just the result.

$\Pr(X|Y) = \Pr(X,Y) / \Pr(Y)$
 $\Pr(X=1 | Y=2) = \Pr(X=1, Y=2) / \Pr(Y=2) = 0.05 / 0.25 = 0.20$

- e) Calculate the conditional probability that a person is ill, given the blood value is between 0 and 20.
- Rewrite the question using complete notation $\Pr(\dots) = \dots$
 - Use Bayes Rule and the previous results

... calculate probability that a person is ill, given the blood value ... $\rightarrow \Pr(Y=2 | X=1) = ?$
 $\Pr(Y=2 | X=1) = \Pr(X=1 | Y=2)\Pr(Y=2) / \Pr(X=1) = 0.2 \times 0.25 / 0.55 = 0.0909$