Technische Hochschule Rosenheim

Chapter 02b – Conditional Probabilities

Lecture A2I2

Prof. Dr. Kai Höfig

Prior Probabilities

- A **Prior Probability** is a probability value that is based on measurements or preexisting knowledge.
  - Example: We measure the velocity of 100 cars and 30% drive too fast. If we now assume that the probability for any single car of driving too fast is 0.3, this is a prior probability.
- A special case of such an assumption is the discrete and equal distribution of probabilities of basic events in a Laplace probability space (indifference principle). Such a **Laplace-Probabilit**y is also an assumption and so a prior probability.
  - Example: The probability of a dice throw (basic event) is 1/6, which is only a assumption. Every dice will have its own distribution of values, depending on the physical precision during production.
- If further probabilities are calculated using prior probabilities, the **Prior Assumption** should hold, for example supported by representative measurements.

Laplace Probabilities

- Let $\Omega = \{\omega_1, \ldots, \omega_n\}$ be a discrete final event space with *n* basic events.

- The probability of basic events is equally distributed $P(\omega_i) = \frac{1}{|\Omega|}$
- The probability of an event A is $P(A) = \frac{|A|}{|\Omega|}$
- Example dice

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

$$P(\omega) = \frac{1}{6}, \forall \omega \in \Omega$$

$$P(gerade\ Zahl) = \frac{|\{2,4,6\}|}{\{|1,2,3,4,5,6|\}} = \frac{1}{2}$$

Basic rules in Laplace event spaces

1. $P(\Omega) = 1$
2. $P(\emptyset) = 0$
3. $P(A \vee B) = P(A) + P(B) - P(A \wedge B)$ for any basic events
4. $P(A \vee B) = P(A) + P(B)$ for pairwise incompatibel events
5. $P(A \wedge B) = P(A) \cdot P(B)$ for independent events A and B
6. $P(A) + P(\neg A) = 1$
7. For $A \subseteq B$ it is $P(A) \leq P(B)$
8. Let $\omega_1, \ldots, \omega_n$ be the basic events. It is $\sum_{i=1}^{n} P(\omega_i) = 1$

Conditional Probability

- For two events A and B, the conditional probability P(A|B) for A under the condition B is

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

- For a finite event space, it is

$$P(A|B) = \frac{P(A \wedge B)}{P(B)} = \frac{\frac{|A \wedge B|}{|\Omega|}}{\frac{|B|}{|\Omega|}} = \frac{|A \wedge B|}{|B|}$$

- If A is independent from B, it is

$$P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{|A| \cdot |B|}{|B|} = P(A)$$

Chain rule

- For any two events A and B it is

$$P(A \wedge B) = P(A|B) \cdot P(B)$$

- For more events it is

$$\begin{aligned}
P(A_1 \wedge \cdots \wedge A_n) &= P(A_1|A_2 \wedge \cdots \wedge A_n) \cdot P(A_2 \wedge \cdots \wedge A_n) \\
&= P(A_1|A_2 \wedge \cdots \wedge A_n) \cdot P(A_2|A_3 \wedge \cdots \wedge A_n) \cdot P(A_3 \wedge \cdots \wedge A_n) \\
& \quad \ldots \\
&= \prod_{k=1}^{n} = P(A_k | \bigwedge_{j=1}^{k-1} A_j)
\end{aligned}$$

Example for Conditional Probabilities (1)

- During speed monitoring, we measure the following and use the relative occurrences of events as prior probabilities

Event	Absolut occurrence	Relative occurrence
Vehicle measured	100	100%
Driver is student (S)	30	30%
Velocity too high (V)	10	10%
Velocity too high and driver is student ($S \wedge V$)	5	5%

- Do students drive too fast more often than the average driver?

$$P(V|S) = \frac{P(V \wedge S)}{P(S)} = \frac{5}{30} \approx 0.17 > P(V) = 0.1$$

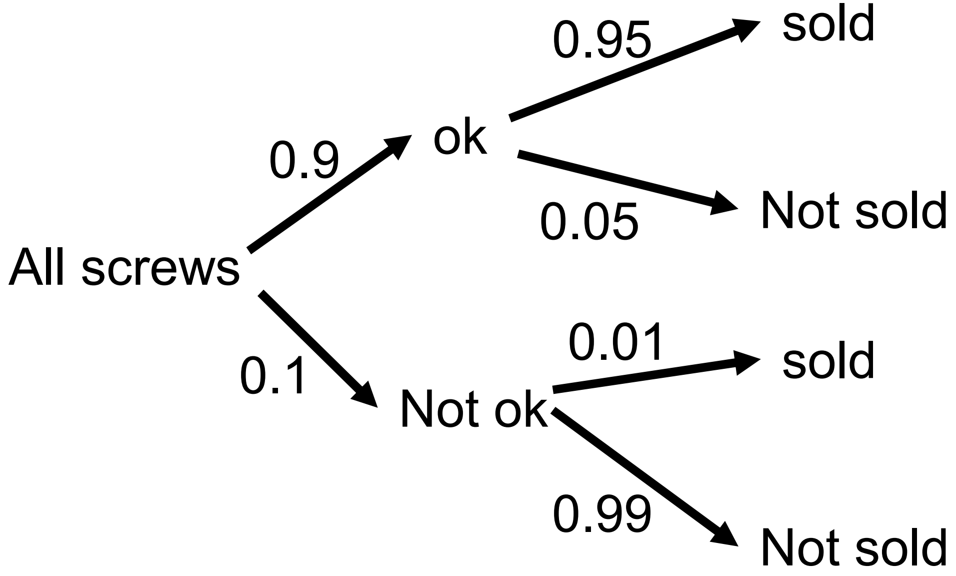
Yes they do, under the condition that the prior probabilities are representative.

Do we know the distribution of our random variables?

	Student	Not a student	Sum
Velocity too high	**0.05**	*0.05*	**0.10**
Velocity not too high	*0.25*	*0.65*	*0.9*
Sum	**0.30**	*0.7*	

Example for Conditional Probabilities (2)

- 90% of produced screws are good, 95 of the good screws are sold, 1% of the bad screws are sold by mistake.



$$P(Screw\ Sold) = 0.9 \cdot 0.95 + 0.1 \cdot 0.01 = 0.856$$

$$P(Screw\ sold\ and\ OK) = 0.9 \cdot 0.95 = 0.855$$

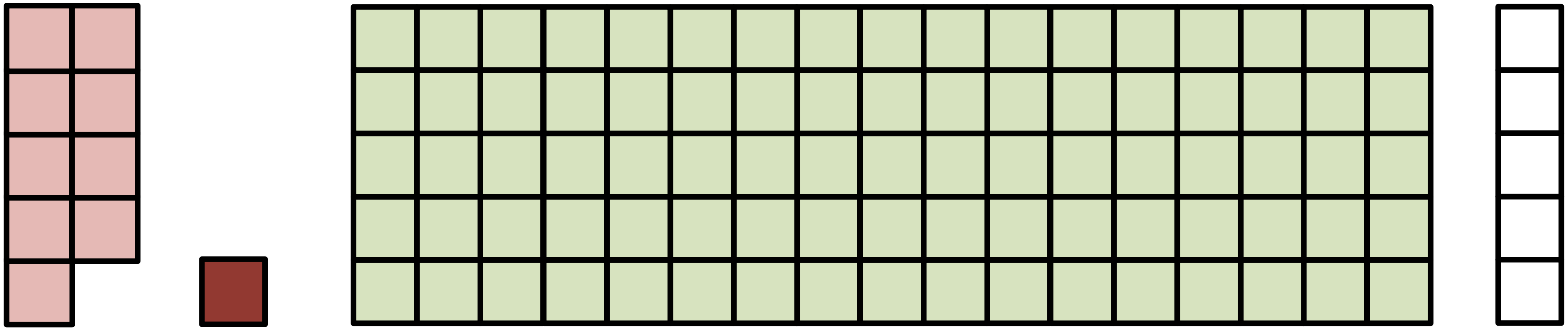
$$P(a\ bought\ screw\ is\ OK) = P(OK|Sold) = \frac{P(S \wedge OK)}{P(S)} = \frac{0.855}{0.856} \approx 0.988$$

Conditional Probability Graphically

90% of produced screws are good, 95 of the good screws are sold, 1% of the bad screws are sold by mistake.

$$P(\neg OK) = 0.10 \qquad P(OK) = 0.90$$

$$P(S|\neg OK) = 0.01 \qquad P(S|OK) = 0.95$$

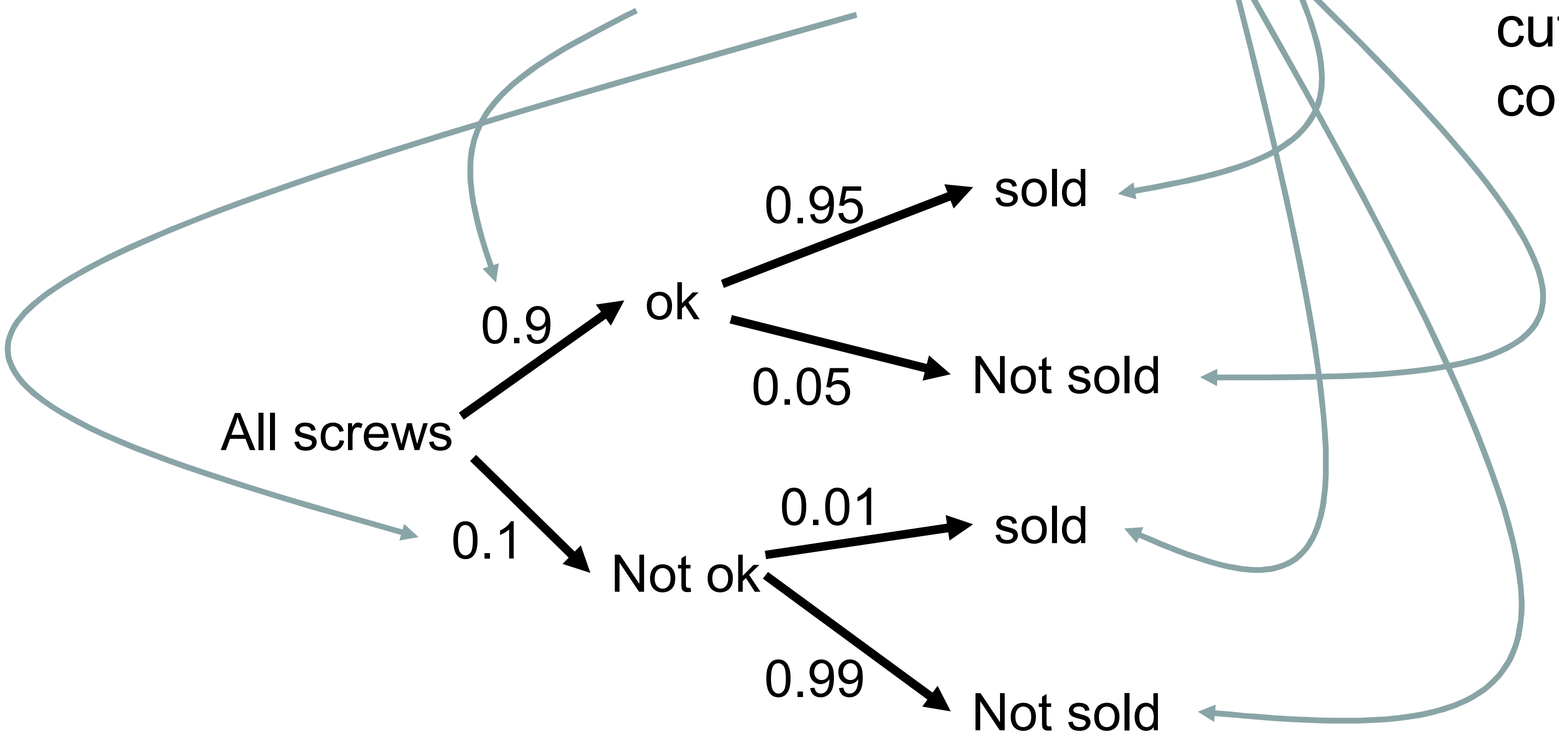


$$P(OK|S) = \frac{P(OK \wedge S)}{P(S)} = \frac{P(S|OK) \cdot P(OK)}{P(S)}$$

$$P(S) = P(OK) \cdot P(S|OK) + P(\neg OK) \cdot P(S|\neg OK)$$

Joint Probability Distribution of the Variables OK and Sold

	OK	Not OK	Sum
Sold	0.855	0.001	0.856
Not Sold	0.045	0.099	0.144
Sum	0.9	0.1	



- The distribution of two or more random variables shows us the probabilities for all the cuts. With these cuts, we can calculate all conditional probabilities.

Marginal Distribution

$$A \Leftrightarrow (A \wedge B) \vee (A \wedge \neg B)$$
$$\Rightarrow P(A) = P((A \wedge B) \vee (A \wedge \neg B))$$
$$= P(A \wedge B) + P(A \wedge \neg B)$$

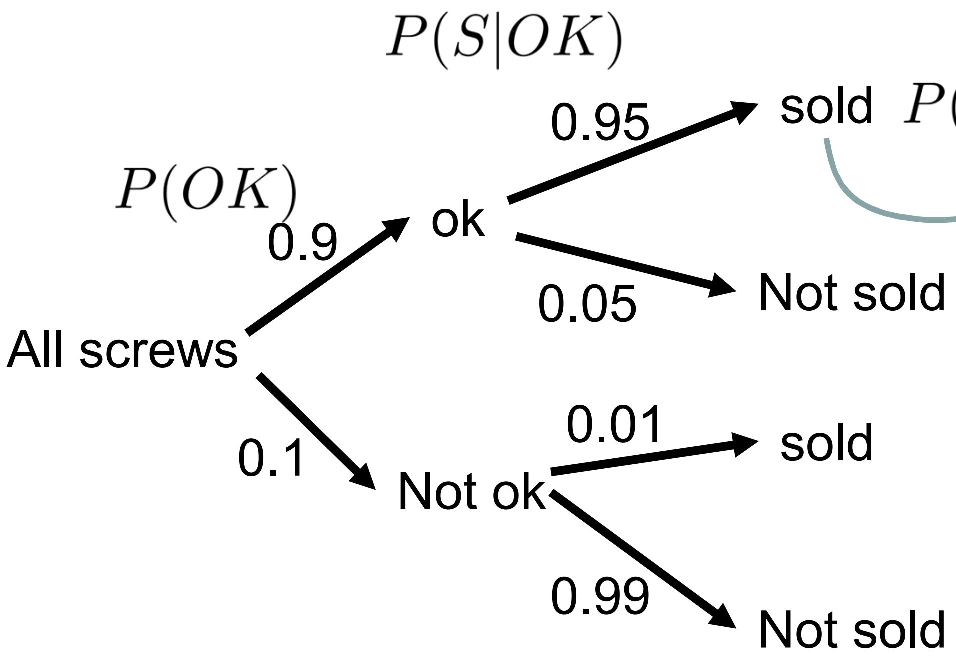
	OK	Not OK	Sum
Sold	0.855	0.001	0.856
Not Sold	0.045	0.099	0.144
Sum	0.9	0.1	

- Marginalization is the process of loosing one dimension and loosing one random variable (in this case B). The resulting probabilities are the marginal distributions of the left random variables.

- We can calculate the marginal probability distribution, if we know the joint probability distribution.

$$P(X_1 = x1, \ldots X_{d-1} = x_{d-1} = \sum_{x_d} P(X_1 = x_1, \ldots X_{d-1} = x_{d-1}, X_d = x_d))$$

Chain Rule



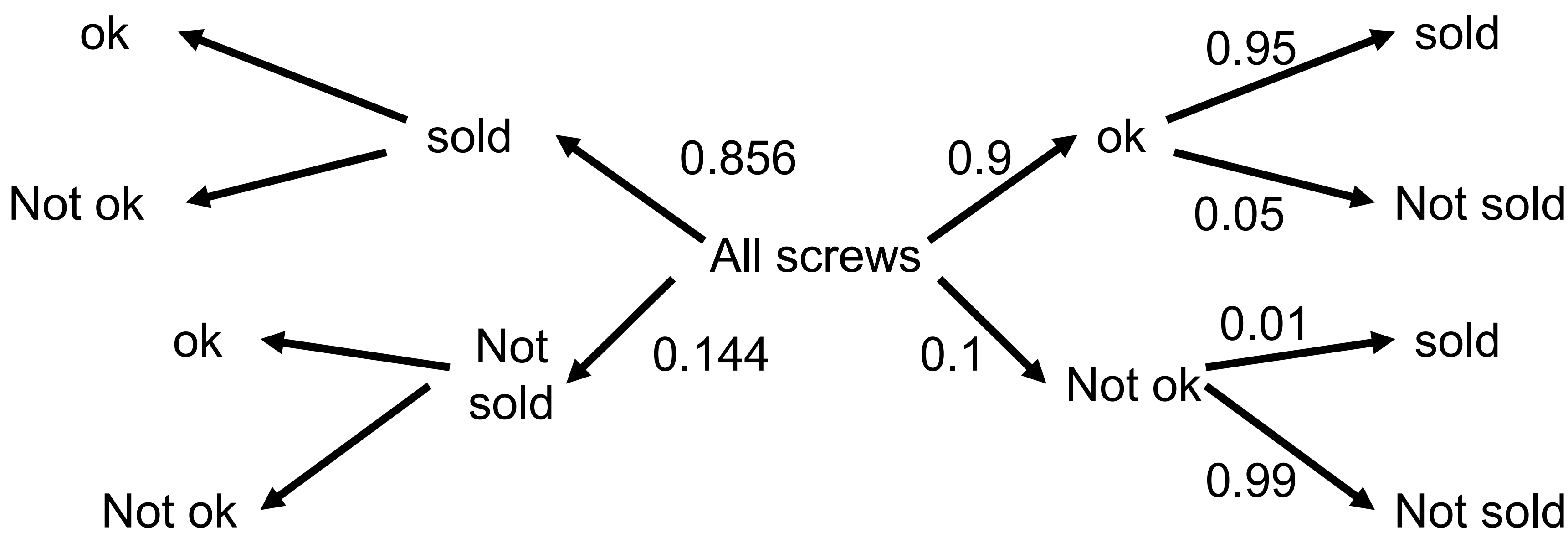
$$P(S \wedge OK) = P(OK) \cdot P(S|OK)$$

	OK	Not OK	Sum
Sold	0.855	0.001	0.856
Not Sold	0.045	0.099	0.144
Sum	0.9	0.1	

- For more than two dimensions, this calculates to

$$P(X_1 \wedge \dots \wedge X_n) = \prod_{i=1}^{n} P(X_i|X_1, \dots, X_{i-1})$$

Bayes' theorem

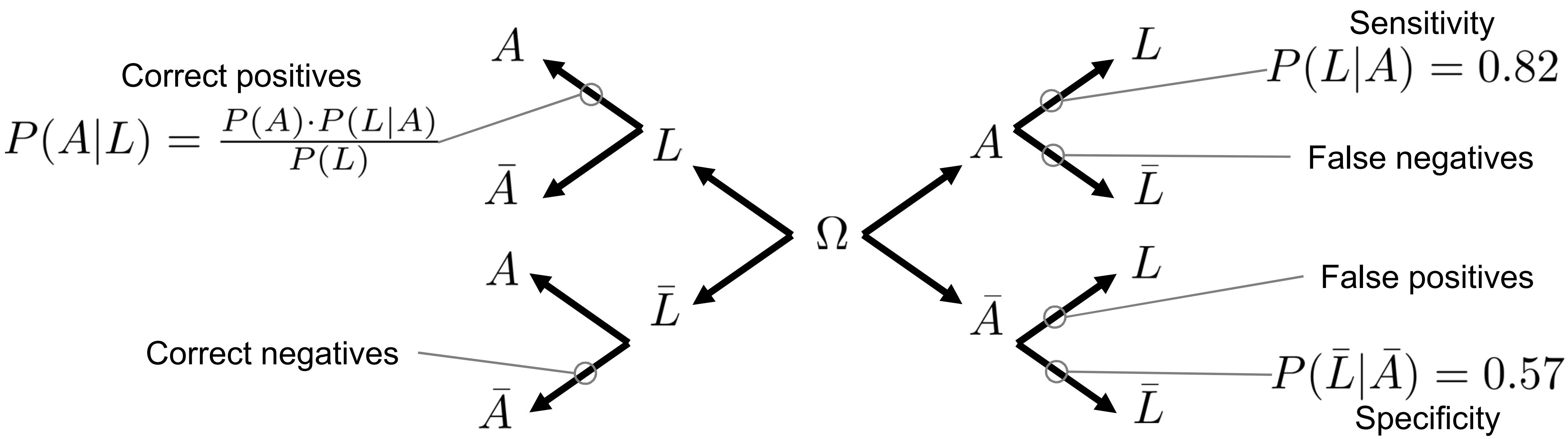


$$
\begin{aligned}
P(OK \wedge S) &= P(S \wedge OK) && (1)\\
\Leftrightarrow P(OK|S) \cdot P(S) &= P(OK) \cdot P(S|OK) && (2)\\
\Leftrightarrow P(OK|S) &= \frac{P(OK) \cdot P(S|OK)}{P(S)} && (3)
\end{aligned}
$$

Appendicitis Example

- The probability of increased white blood cells when a person has appendicitis is $P(L|A) = 0.82$ (Sensitivity of the test)

- The probability of normal white blood cell concentration if a person has no appendicitis is $P(\bar{L}|\bar{A}) = 0.57$ (Specificity of the test)



Depending on demographic properties such as the probability for appendicitis and other causes for a high white blood cell concentration

Can be measured during experiments, universal value for the test