Pravděpodobnost a matematická statistika

Doc. RNDr. Gejza Dohnal, CSc.
dohnal@nipax.cz

Úvod do teorie pravděpodobnosti

- Náhoda a pravděpodobnost,
- náhodný jev, náhodná veličina
- rozdělení pravděpodobnosti a jeho souvislost s histogramem
- pravidla pro počítání s pravděpodobností
- podmíněná pravděpodobnost
- závislost náhodných veličin
- využití závislosti při stanovení pravděpodobnosti - věta o úplné pravděpodobnosti a Bayesova věta

Úvod do teorie pravděpodobnosti

Náhodný pokus $=$ činnost, která může skončit různým výsledkem a my dopředu nevíme kterým

Náhodný jev $=$ výsledek náhodného pokusu, o kterém dopředu nevíme, zda nastane či ne

Pravděpodobnostní prostor $(\Omega, \mathcal{F}, P)$
(Kolmogorov, 1933)

Ω - množina elemetárních náhodných jevů

$\mathcal{F}$ - jevové pole (σ-algebra náhodných jevů)

(a) $\emptyset, \Omega \in \mathcal{F}$
(b) pokud $A \in \mathcal{F}$, potom platí $A^C \in \mathcal{F}$
(c) pokud $A_1, A_2, \dots \in \mathcal{F}$, potom platí $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

P - pravděpodobnostní míra na $\mathcal{F}$

Úvod do teorie pravděpodobnosti

Definice pravděpodobnosti

1. Klasická definice pravděpodobnosti: $P(A) = \dfrac{\mu(A)}{\mu(\Omega)}$

   $\mu(A) =$ počet elementárních jevů v jevu A

   $\mu(A) =$ "míra" jevu A

2. Axiomatická definice pravděpodobnosti

   (a) $\forall A \in \mathcal{F} : \quad P(A) \in \langle 0, 1 \rangle$

   (b) $P(\emptyset) = 0, \quad P(\Omega) = 1$

   (c) jsou-li $A_1, A_2, \dots \in \mathcal{F}$, po dvou disjunktní, potom

   $$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$$

3. Statistická definice pravděpodobnosti: $\lim_{j \to \infty} \dfrac{n_j^A}{n_j}$

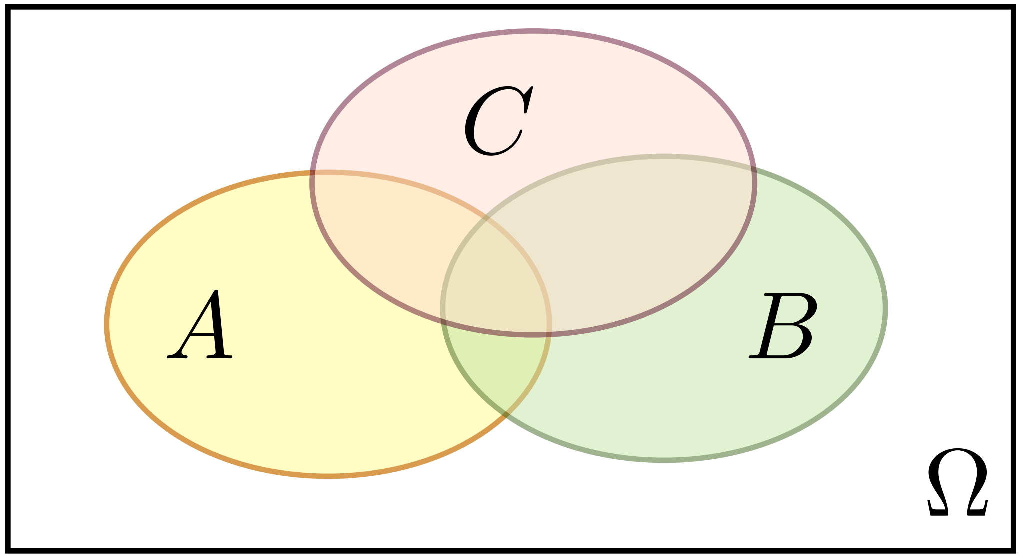
Úvod do teorie pravděpodobnosti

Pravidla pro počítání s pravděpodobností

1) $P(A) \in \langle 0, 1 \rangle$

2) $P(\emptyset) = 0, \quad P(\Omega) = 1$

3) $P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P\left(A_i\right)$



4) $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

5) $P(A - B) = P(A) - P(A \cap B)$

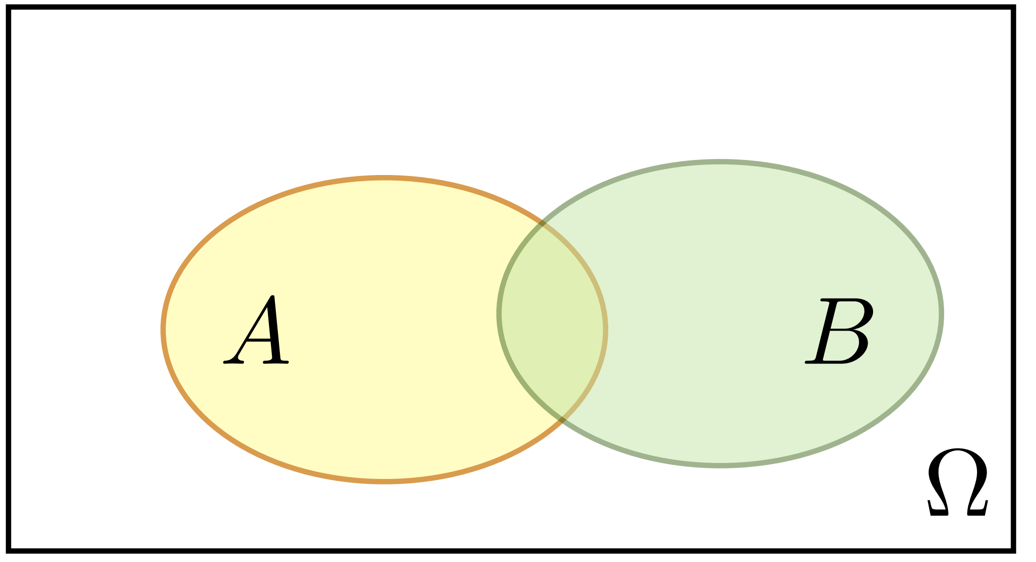
6) $P(A^C) = 1 - P(A)$

$$\boxed{P(A \cap B) = ?}$$

Podmíněná pravděpodobnost

$$P(A \cap B) = P(A).P(B|A)$$

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



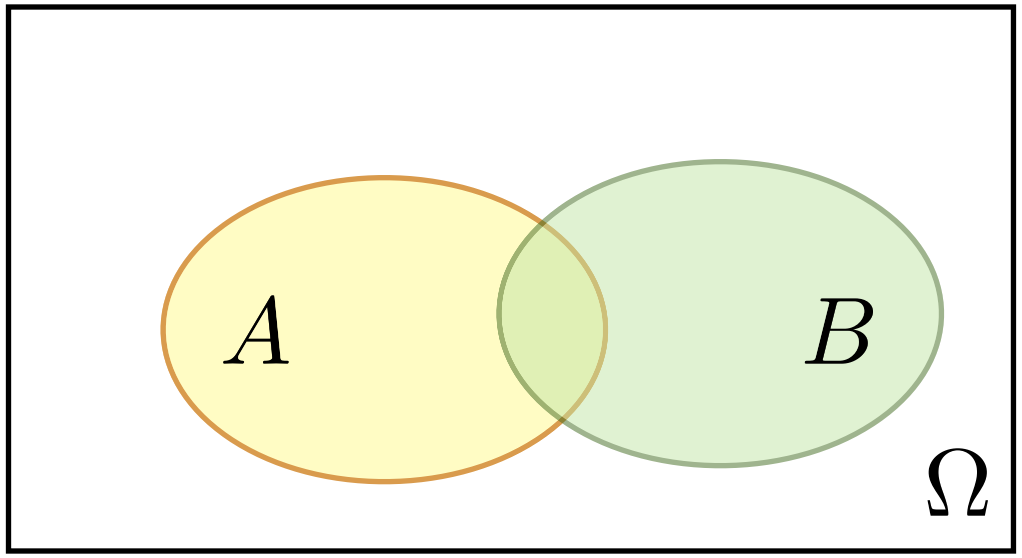
Příklad:

Je známo, že v dlouhodobém průměru je mezi 1000 dodaných komponent 2,34% vadných výrobků výrobce A, 1,08% vadných výrobků výrobce B, 65,97% bezvadných výrobků výrobce A a zbytek (30,6%) jsou bezvadné výrobky od výrobce B. Lze považovat jev, že výrobek je vadný, za stochasticky závislý na výrobci?

Podmíněná pravděpodobnost

$$P(A \cap B) = P(A).P(B|A)$$

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



Příklad:

Jevy A a V jsou stochasticky nezávislé právě když

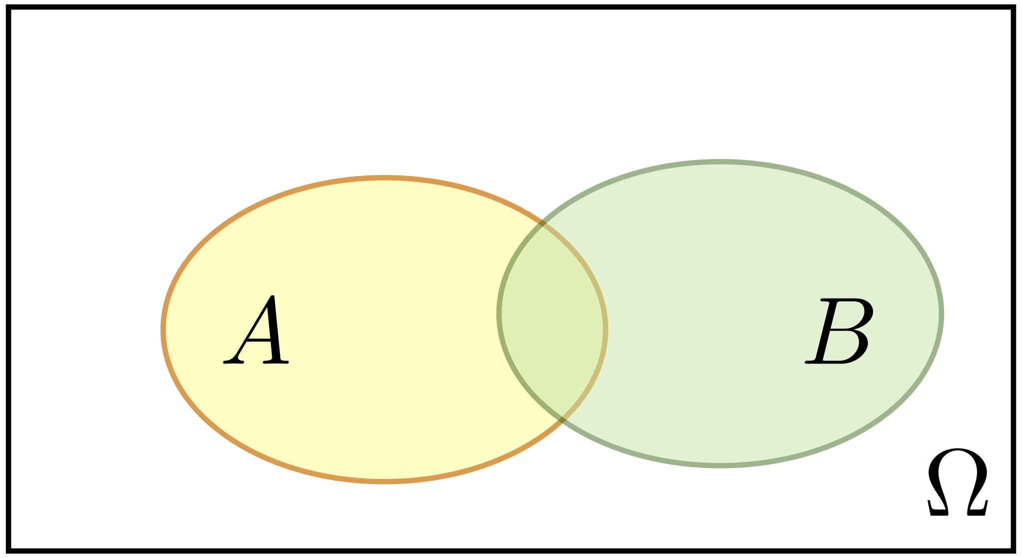
$$P(V|A) = P(V)$$
$$P(A|V) = P(A)$$
$$P(A \cap V) = P(A).P(V)$$

	A	B	celkem
vada (V)	2,34	1,08	3,42
bez vady	65,98	30,6	96,58
celkem	68,32	31,68	100

Podmíněná pravděpodobnost

$$P(A \cap B) = P(A).P(B|A)$$

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



Příklad:

Jev, že výrobek je vadný, lze považovat za stochasticky nezávislý na výrobci.

$$2{,}34 = 3{,}42 \,.\, 68{,}32$$
$$1{,}08 = 3{,}42 \,.\, 31{,}68$$
$$65{,}98 = 96{,}58 \,.\, 68{,}32$$
$$30{,}60 = 96{,}58 \,.\, 31{,}68$$

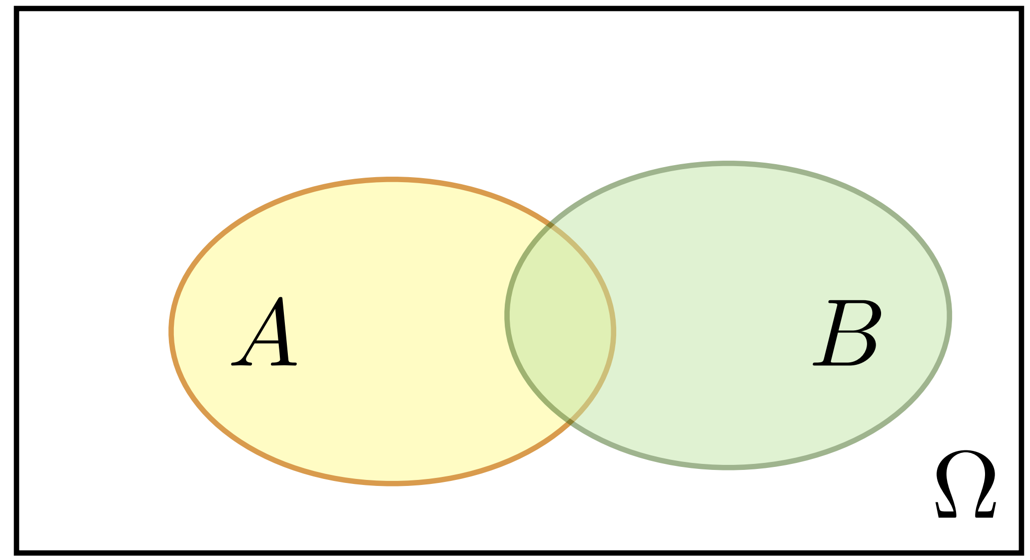
o.K.

	A	B	celkem
vada (V)	2,34	1,08	3,42
bez vady	65,98	30,6	96,58
celkem	68,32	31,68	100

Podmíněná pravděpodobnost

$$P(A \cap B) = P(A).P(B|A)$$

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



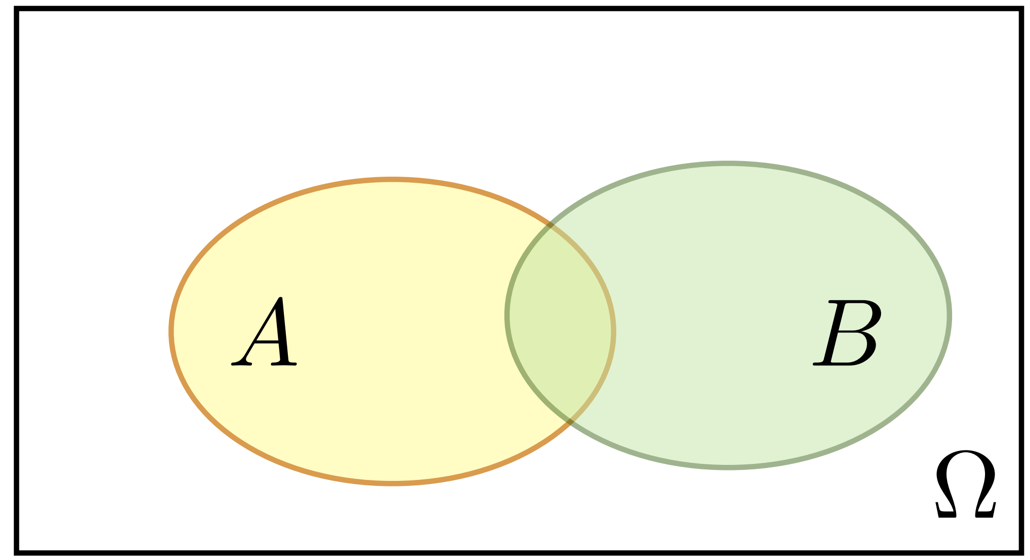
Příklad:

Z šetření obchodního řetězce vyplývá, že 34% zákazníků nakupuje pivo Pardál a současně dámské kalhotky. 28% zákazníků nakupuje dámské kalhotky, ale nenakupuje pivo Pardál, 19% nakupuje Pardála a nekupuje dámské kalhotky. Ostatních 19% nekupuje ani jednu z komodit. Lze považovat nákup piva Pardál stochasticky nezávislý na nákupu dámských kalhotek?

Podmíněná pravděpodobnost

$$P(A \cap B) = P(A).P(B|A)$$

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



Příklad:

$$0{,}62\ .\ 0{,}53 = 0{,}3286$$

$$P(K) = 0{,}62$$

$$P(K|P) = 0{,}42/0{,}62$$

$$= 0{,}677 > P(K)$$

Nákup dámských kalhotek je stochasticky závislý na nákupu piva Pardál

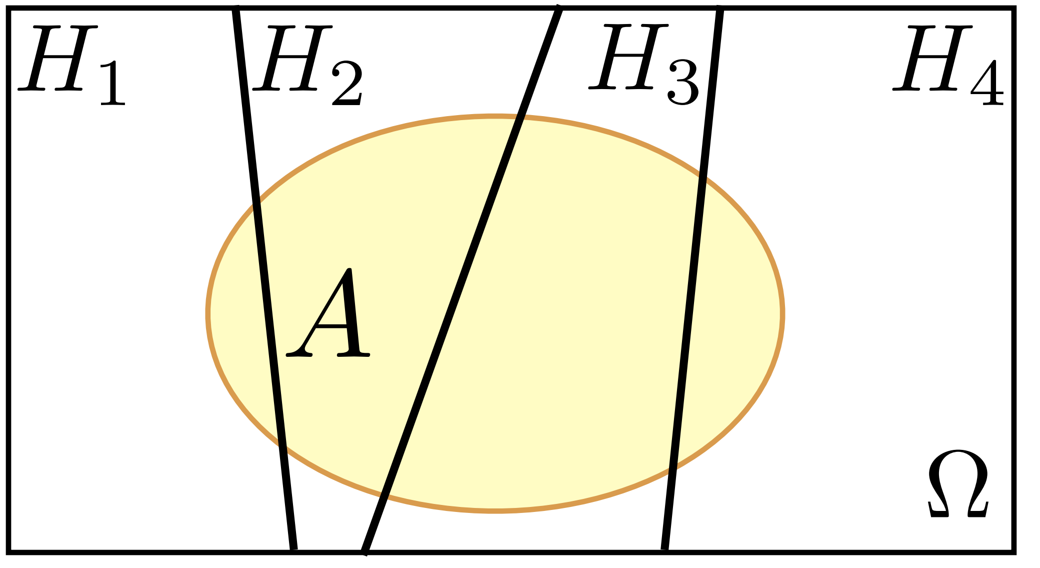
	P	P^C	celkem
K	0,42	0,2	0,62
K^C	0,19	0,19	0,38
celkem	0,53	0,47	1

Věta o úplné pravděpodobnosti

A je náhodný jev, $\{H_1, H_2, H_3, H_4\}$ je úplné pokrytí Ω

$$H_i \cap H_j = \emptyset, \quad H_1 \cup H_2 \cup H_3 \cup H_4 = \Omega$$

$$\boxed{\begin{aligned} P(A) = P(A|H_1).P(H_1) \\ +P(A|H_2).P(H_2) \\ +P(A|H_3).P(H_3) \\ +P(A|H_4).P(H_4) \end{aligned}}$$



Příklad:

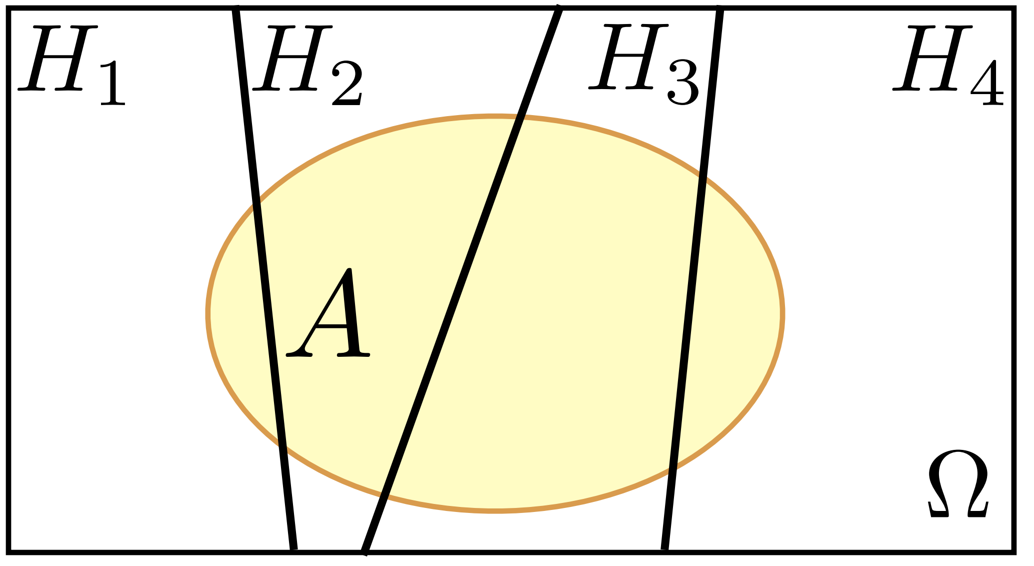
Na trhu jsou výrobky od čtyř výrobců v pořadí A, B, C a D v poměru 1:2:4:5. Zmetkovitost je u těchto výrobců po řadě 0,5%, 0,8%, 0,3% a 0,3%. Jaká je pravděpodobnost, že náhodně vybraný výrobek na trhu bude vadný? A bude-li vadný, jaká je pravděpodobnost, že byl vyroben výrobcem A?

Věta o úplné pravděpodobnosti

A je náhodný jev, $\{H_1, H_2, H_3, H_4\}$ je úplné pokrytí Ω

$$H_i \cap H_j = \emptyset, \quad H_1 \cup H_2 \cup H_3 \cup H_4 = \Omega$$

$$\boxed{\begin{aligned} P(A) = {} & P(A|H_1).P(H_1) \\ & +P(A|H_2).P(H_2) \\ & +P(A|H_3).P(H_3) \\ & +P(A|H_4).P(H_4) \end{aligned}}$$



Příklad:

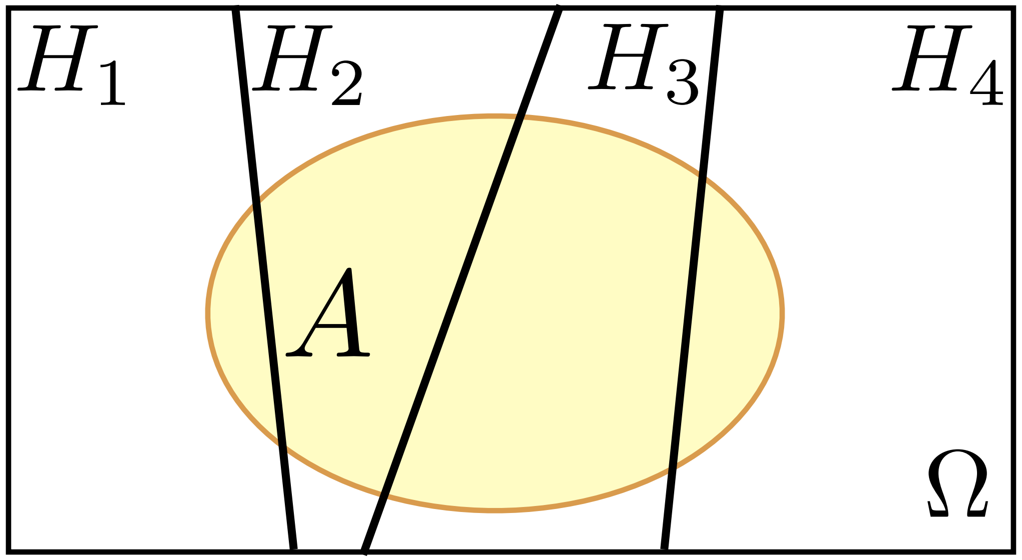
$$P(H_1) = 1/12, \; P(A|H_1) = 0{,}005$$
$$P(H_2) = 2/12, \; P(A|H_2) = 0{,}008$$
$$P(H_3) = 4/12, \; P(A|H_3) = 0{,}003$$
$$P(H_4) = 5/12, \; P(A|H_4) = 0{,}003$$
$$\boxed{\mathrm{P}(A) = (5 + 16 + 12 + 15)/12000 = 0{,}004}$$

Bayesova věta

A je náhodný jev, $\{H_1, H_2, H_3, H_4\}$ je úplné pokrytí Ω

$$H_i \cap H_j = \emptyset, \quad H_1 \cup H_2 \cup H_3 \cup H_4 = \Omega$$

$$P(H_1|A) = \frac{P(A|H_1).P(H_1)}{P(A)}$$

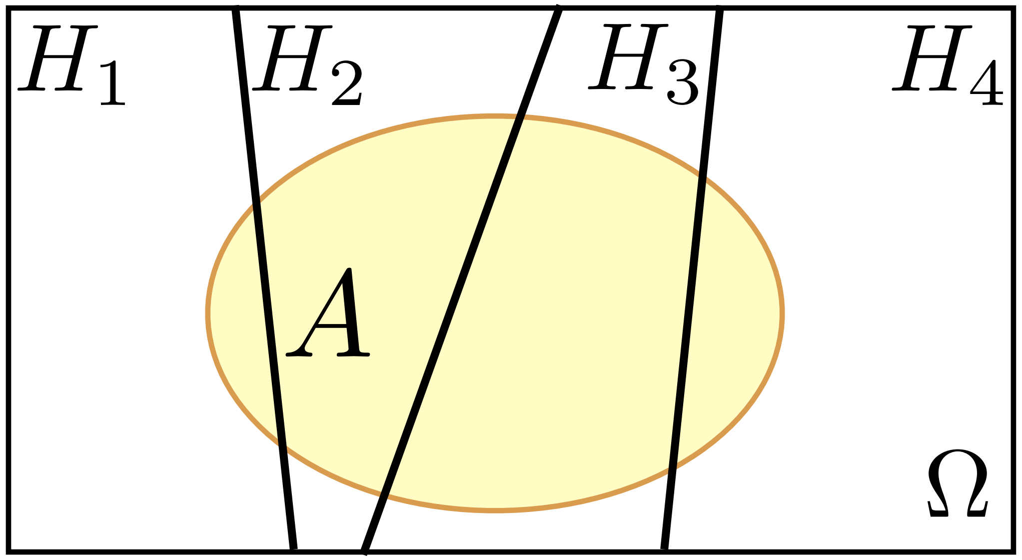


Bayesova věta

A je náhodný jev, $\{H_1, H_2, H_3, H_4\}$ je úplné pokrytí Ω

$$H_i \cap H_j = \emptyset, \quad H_1 \cup H_2 \cup H_3 \cup H_4 = \Omega$$

$$P(H_1|A) = \frac{P(A|H_1).P(H_1)}{\sum_{i=1}^{4} P(A|H_i).P(H_i)}$$



Příklad:

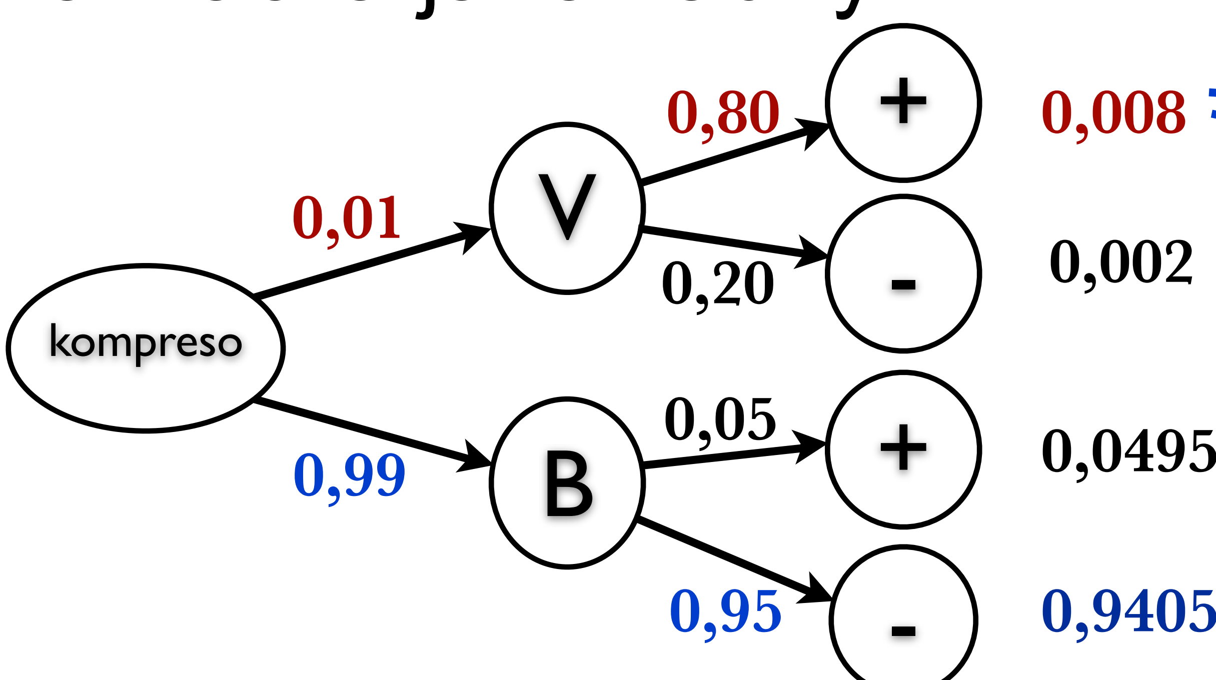
$$P(H_1) = 0{,}08333, \quad P(A|H_1) = 0{,}005$$
$$P(H_2) = 0{,}16667, \quad P(A|H_2) = 0{,}008$$
$$P(H_3) = 0{,}33333, \quad P(A|H_3) = 0{,}003$$
$$P(H_4) = 0{,}41667, \quad P(A|H_4) = 0{,}003$$

$$\boxed{\mathrm{P}(H_1|A) = 0{,}005.0{,}08333/0{,}004 = 0{,}10417}$$

Bayesova věta

Příklad:

Tělo kompresoru musí být 100% hermeticky uzavřeno. V průměru jeden ze sta kompresorů trochu netěsní a je třeba jej rozložit a znovu složit. Zkouška těsnosti odhalí vadný výrobek s pravděpodobností 0,8. Naproti tomu, s pravděpodobností 0,05 označí jako vadný bezvadný výrobek. Jaká je pravděpodobnost vady u výrobku, který zkouška označila jako vadný?

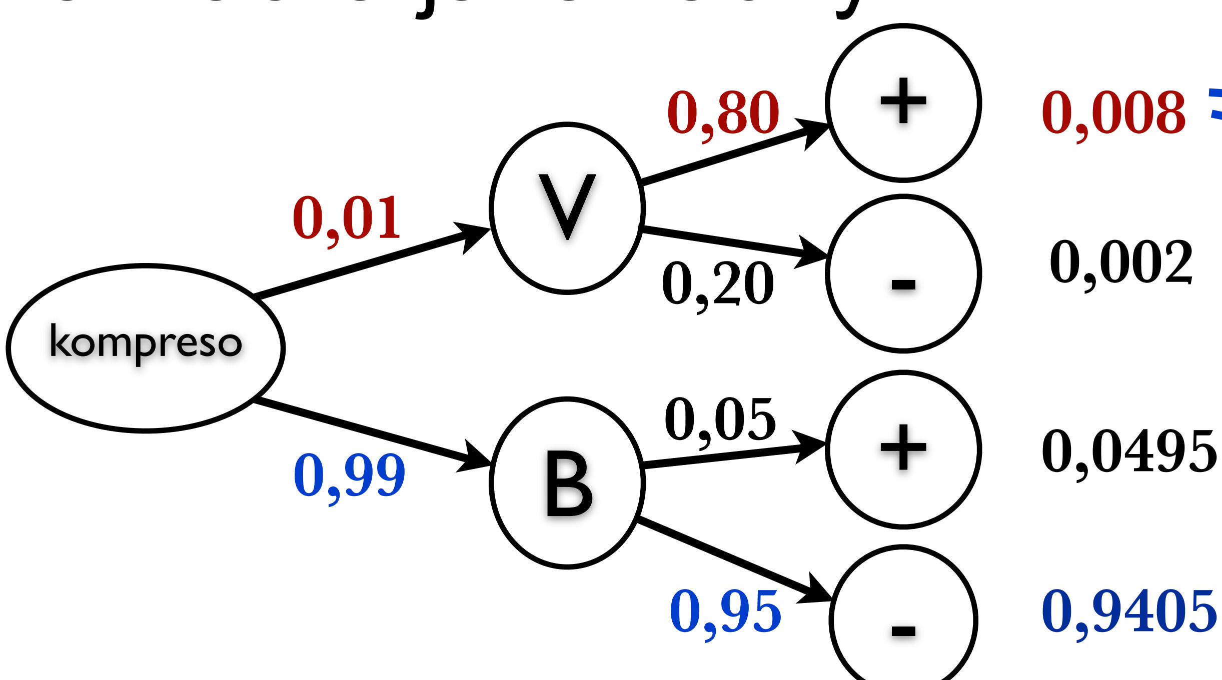


$$P(V|+) = \frac{0{,}008}{0{,}008+0{,}0495}$$

Bayesova věta

Příklad:

Tělo kompresoru musí být 100% hermeticky uzavřeno. V průměru jeden ze sta kompresorů trochu netěsní a je třeba jej rozložit a znovu složit. Zkouška těsnosti odhalí vadný výrobek s pravděpodobností 0,8. Naproti tomu, s pravděpodobností 0,05 označí jako vadný bezvadný výrobek. Jaká je pravděpodobnost vady u výrobku, který zkouška označila jako vadný?

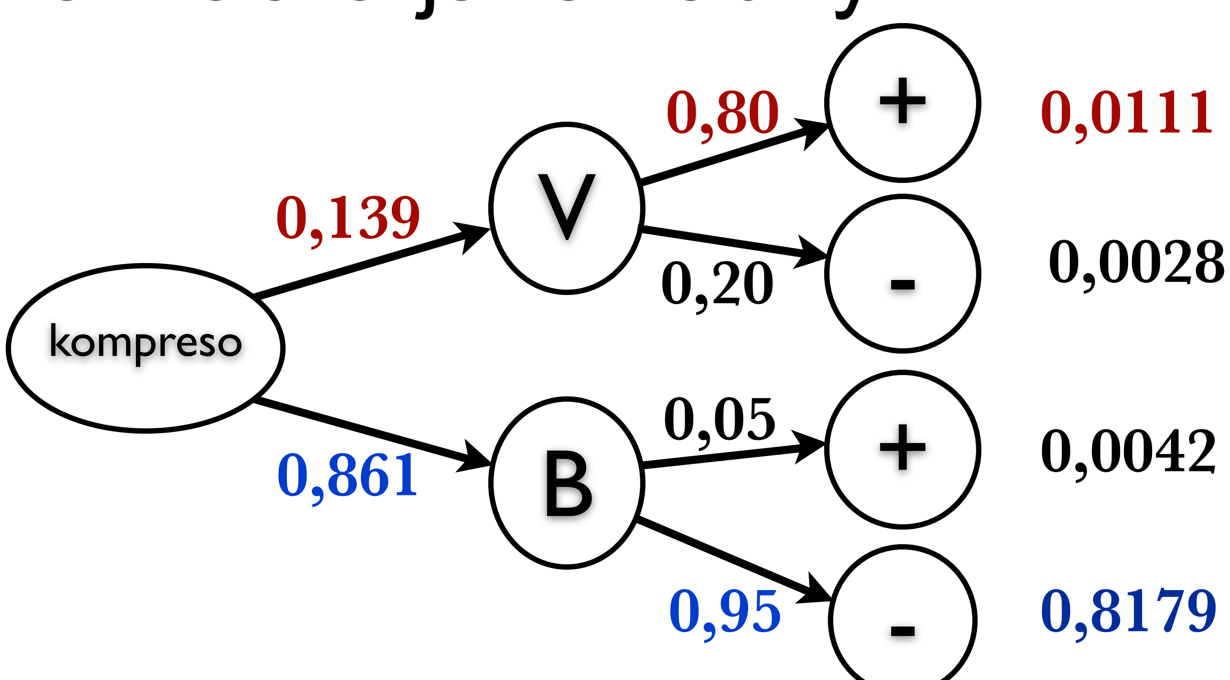


$$P(V|+) = \frac{0{,}008}{0{,}008+0{,}0495} = 0{,}1391$$

Bayesova věta

Příklad:

Tělo kompresoru musí být 100% hermeticky uzavřeno. V průměru jeden ze sta kompresorů trochu netěsní a je třeba jej rozložit a znovu složit. Zkouška těsnosti odhalí vadný výrobek s pravděpodobností 0,8. Naproti tomu, s pravděpodobností 0,05 označí jako vadný bezvadný výrobek. Jaká je pravděpodobnost vady u výrobku, který zkouška označila jako vadný?



$$P_2(V|+) = \frac{0{,}0111}{0{,}0111+0{,}0042}$$

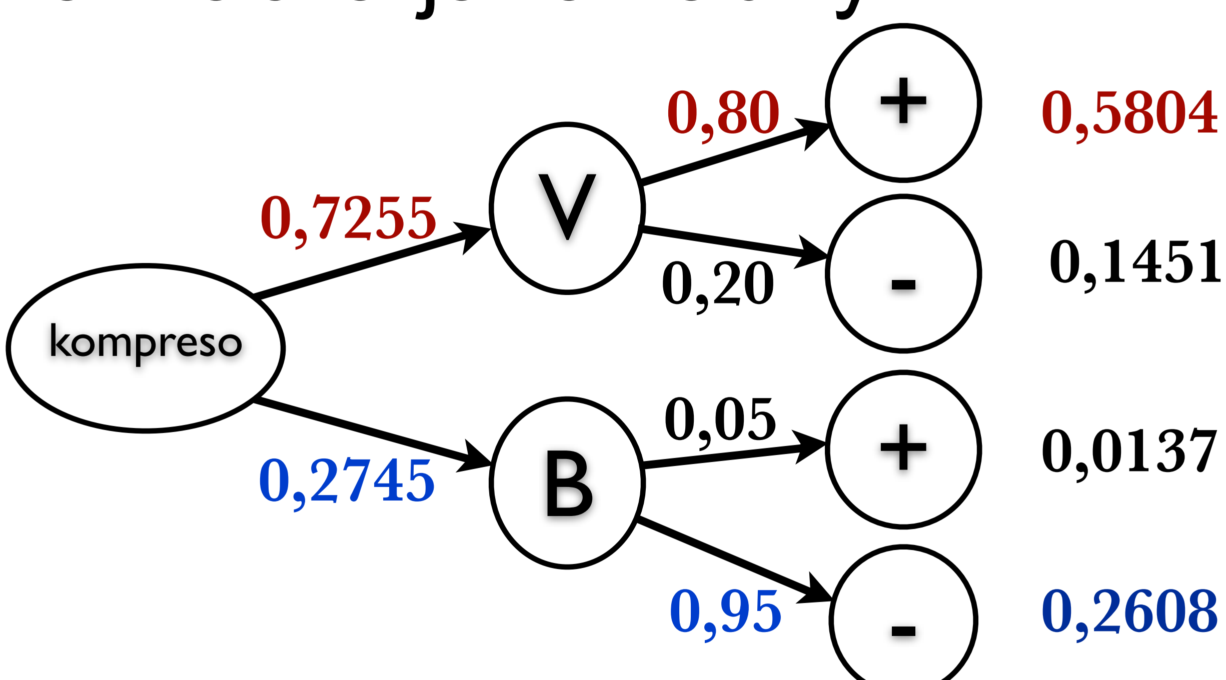
$$\boxed{= 0{,}7255}$$

Při opakovaném testu

Bayesova věta

Příklad:

Tělo kompresoru musí být 100% hermeticky uzavřeno. V průměru jeden ze sta kompresorů trochu netěsní a je třeba jej rozložit a znovu složit. Zkouška těsnosti odhalí vadný výrobek s pravděpodobností 0,8. Naproti tomu, s pravděpodobností 0,05 označí jako vadný bezvadný výrobek. Jaká je pravděpodobnost vady u výrobku, který zkouška označila jako vadný?



$$P_3(V|+) = \frac{0{,}5804}{0{,}5804+0{,}0137}$$

$$\boxed{= 0{,}977}$$

Při opakovaném testu