

Zákon celkové pravděpodobnosti

- úplný systém disjunktálních podmnožin $\{A_i\}$

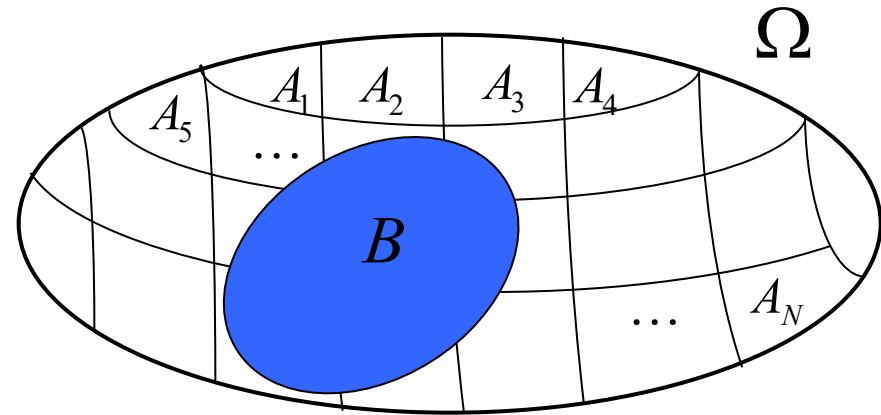
$$\bigcup_{i=1}^N A_i = \Omega$$

$$A_i \cap A_j = \emptyset \quad i \neq j$$

$$B = \bigcup_{i=1}^N (B \cap A_i)$$

$$P(B) = \sum_{i=1}^N P(B \cap A_i) = \sum_{i=1}^N P(B|A_i)P(A_i)$$

$$P(B) = \sum_{i=1}^N P(B|A_i)P(A_i)$$



- Bayesův teorém

$$P(A_i|B) = \frac{P(B|A_i)P(A_i)}{P(B)} = \frac{P(B|A_i)P(A_i)}{\sum_{i=1}^N P(B|A_i)P(A_i)}$$

Podmíněná pravděpodobnost: příklad – dveře s pokladem



Podmíněná pravděpodobnost: příklad – dveře s pokladem

$$\downarrow P(A_C) = \frac{1}{3}$$



M – hostitel otevře modré dveře $P(M) = \frac{1}{2}$
(pokud má tu možnost vybírá dveře náhodně)

A_C – poklad je za červenými dveřmi $P(M|A_C) = \frac{1}{2}$

A_M – poklad je za modrými dveřmi $P(M|A_M) = 0$

A_Z – poklad je za žlutými dveřmi $P(M|A_Z) = 1$

$$P(A_C|M) = \frac{P(M|A_C)P(A_C)}{P(M)} = \frac{\frac{1}{2} \frac{1}{3}}{\frac{1}{2}} = \frac{1}{3}$$

$$P(A_M|M) = \frac{P(M|A_M)P(A_M)}{P(M)} = \frac{0 \times \frac{1}{3}}{\frac{1}{2}} = 0$$

$$P(A_Z|M) = \frac{P(M|A_Z)P(A_Z)}{P(M)} = \frac{1 \times \frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

Podmíněná pravděpodobnost: příklad – Máte děťátko?



prostor jevů: $\Omega = \{(t,+), (t,-), (n,+), (n,-)\}$

$T = \{(t,+), (t,-)\}$ – je těhotná

$$P(T) = 0.008$$

$N = \{(n,+), (n,-)\}$ – není těhotná

$$P(N) = 0.992$$

$R_+ = \{(t,+), (n,+)\}$ – test říká ano

$$P(R_+ | T)$$

$R_- = \{(t,-), (n,-)\}$ – test říká ne

index selhání

10 žen za 1 rok ze 100 žen

$$\Rightarrow P(T \text{ za 1 rok}) = 0.1$$

$$\Rightarrow P(T \text{ za 1 měsíc}) = 0.1/12 \cong 0.008$$

$$P(T | R_+) = \frac{P(R_+ | T)P(T)}{P(R_+ | T)P(T) + P(R_+ | N)P(N)}$$

Podmíněná pravděpodobnost: příklad – Máte děťátko?



Instant-View® Pregnancy Urine Test



Method	Lateral flow test
Detection level	25 mIU/ml
Accuracy	> 99%
Sample	Urine
Read time	2-5 minutes
Storage	Room temperature
Shelf life	24 months from date of manufacture
CLIA Classification	Waived



NEW Early Detection Ultra Pregnancy test strip Pack.

- **Convenient:** Mailed to your Home.
- **Cost:** Lowest Internet price guarantee
- **Easy:** One Step System.
- **Early:** Detects **10 mIU/ml** of hCG.
- Can work up to 3 days earlier than standard tests.)
- **Fast:** Reliable results in three minutes.
- **Accurate:** **Greater than 99% accuracy, sensitivity and specificity in clinical trials.**
- **European:** Includes multi-lingual Instructions.
- **Delivery:** Same day dispatch **1st Class UK delivery.**
- Insert the number of packs you wish to purchase in the box
- next to the selected item and "add to Basket"

Podmíněná pravděpodobnost: příklad – Máte děťátko?



prostor jevů: $\Omega = \{(t,+), (t,-), (n,+), (n,-)\}$

index selhání

10 žen za 1 rok ze 100 žen

$$\Rightarrow P(T \text{ za 1 rok}) = 0.1$$

$$\Rightarrow P(T \text{ za 1 měsíc}) = 0.1/12 \cong 0.008$$

$T = \{(t,+), (t,-)\}$ – je těhotná

$$P(T) = 0.008$$

$N = \{(n,+), (n,-)\}$ – není těhotná

$$P(N) = 0.992$$

$R_+ = \{(t,+), (n,+)\}$ – test říká ano

$$P(R_+ | T) = 0.995$$

$R_- = \{(t,-), (n,-)\}$ – test říká ne

$$P(R_+ | N) = 0.01$$

$$P(R_+ | N) = 0.001$$

$$P(T | R_+) = 0.89$$

$$P(T | R_+) = \frac{P(R_+ | T)P(T)}{P(R_+ | T)P(T) + P(R_+ | N)P(N)}$$

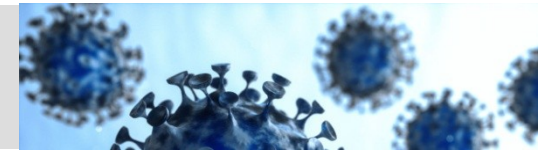
$$P(T | R_+) = 0.45$$



Parametry testu

- senzitivita (citlivost): $P(R_+|T)$
- specificita (síla testu): $P(R_-|N) = 1 - P(R_+|N)$
- věrohodnostní poměr: $\frac{P(R_+|T)}{P(R_+|N)} = \frac{\text{senzitivita}}{1 - \text{specificita}}$

Covid 19 test



- počet aktivních případů v ČR: 168 405
- $$P(T | R_+) = \frac{P(R_+ | T)P(T)}{P(R_+ | T)P(T) + P(R_+ | N)P(N)}$$

- počet obyvatel ČR: 10.71 mil.

nakažených je víc *mám příznaky*

- pravděpodobnost, že jste nakažení: $P(T) = 0.016 \longrightarrow P(T) = 0.032$ $P(T) = 0.50$

- PCR test:

- senzitivita $P(R_+ | T) = 0.98$
- specificita $P(R_- | N) = 0.99$
- věrohodnostní poměr = 98

$$P(T | R_+) = 0.61 \quad P(T | R_+) = 0.76 \quad P(T | R_+) = 0.99$$

$$P(T | R_-) = 0.0003 \quad P(T | R_-) = 0.0007 \quad P(T | R_-) = 0.02$$

- antigenní test:

- senzitivita $P(R_+ | T) = 0.90$
- specificita $P(R_- | N) = 0.97$
- věrohodnostní poměr = 30

$$P(T | R_+) = 0.33 \quad P(T | R_+) = 0.50 \quad P(T | R_+) = 0.97$$

$$P(T | R_-) = 0.002 \quad P(T | R_-) = 0.003 \quad P(T | R_-) = 0.09$$

Podmíněná pravděpodobnost: příklad – Spam filter

bogofilter

$$P(\text{spam}|\text{word}) = \frac{P(\text{word}|\text{spam})P(\text{spam})}{P(\text{word})} = \frac{P(\text{word}|\text{spam})P(\text{spam})}{P(\text{word}|\text{spam})P(\text{spam}) + P(\text{word}|\text{nespam})P(\text{nespam})}$$

$$P\left(\text{spam} \left| \sum_{i=1}^N w_i \right.\right) = \frac{P\left(\sum_{i=1}^N w_i | \text{spam}\right)P(\text{spam})}{P\left(\sum_{i=1}^N w_i | \text{spam}\right)P(\text{spam}) + P\left(\sum_{i=1}^N w_i | \text{nespam}\right)P(\text{nespam})} = \frac{P\left(\sum_{i=1}^N w_i | \text{spam}\right)}{P\left(\sum_{i=1}^N w_i | \text{spam}\right) + P\left(\sum_{i=1}^N w_i | \text{nespam}\right)}$$

1. učení: stanovení pravděpodobností $P(w_i|\text{spam})$ $P(w_i|\text{nespam})$

2. filtrování

$$P(\text{Nigeria}|\text{spam}) = 0.90 \quad P(\text{Nigeria}|\text{nespam}) = 0.01$$

$$P(\text{love}|\text{spam}) = 0.60 \quad P(\text{love}|\text{nespam}) = 0.40$$

$$P(\text{discount}|\text{spam}) = 0.85 \quad P(\text{discount}|\text{nespam}) = 0.10$$

$$P(\text{physics}|\text{spam}) = 0.10 \quad P(\text{physics}|\text{nespam}) = 0.80$$

$$P(\text{spam}) = P(\text{nespam}) = 0.50$$

$$P(\text{spam}) = P(\text{nespam})$$

$$P\left(\sum_{i=1}^N w_i | \text{spam}\right) = \prod_{i=1}^N P(w_i | \text{spam})$$

- Email 1: Nigeria, discount

$$P(\text{spam}|\text{Nigeria} + \text{discount}) = 0.999$$

- Email 2: Nigeria, physics

$$P(\text{spam}|\text{Nigeria} + \text{physics}) = 0.918$$

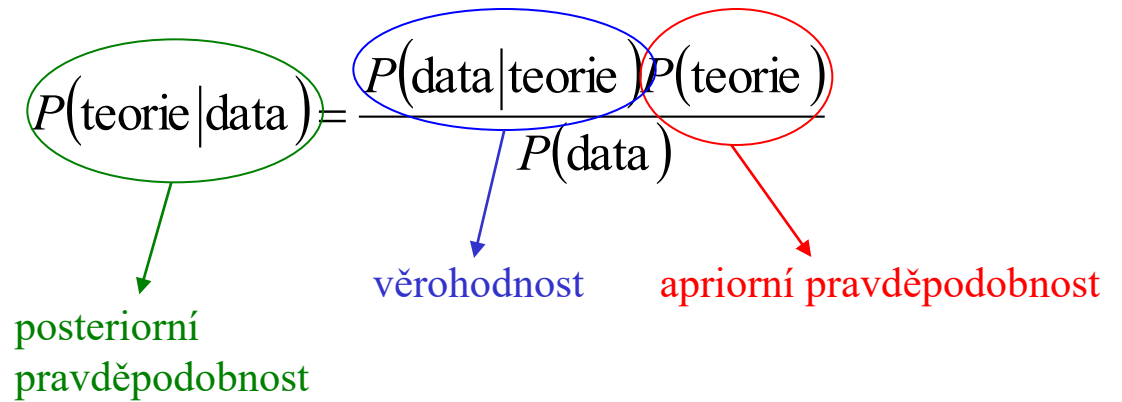
- Email 3: love, physics

$$P(\text{spam}|\text{love} + \text{physics}) = 0.158$$

Podmíněná pravděpodobnost: Bayesův teorém

Bayesův teorém

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



‘update’ víry v platnost teorie po provedení experimentu

$$P(T|D_1 \wedge D_2 \dots \wedge D_N) = P(T) \prod_{i=1}^N \frac{P(D_i|T)}{P(D_i)}$$

Podmíněná pravděpodobnost: příklad - obchodníci z deštěm



$$P(T|D_1) = \frac{P(D_1|T)P(T)}{P(D_1|T)P(T) + P(D_1|\bar{T})P(\bar{T})}$$

$$P(D_1|T) = 1$$

prší 20 dní v roce

$$P(D_1|\bar{T}) = \frac{20}{365} = 0.055$$

$$P(T) = 0.5$$

$$P(T|D_1) = \frac{1 \times 0.5}{1 \times 0.5 + 0.055 \times 0.5} = 0.948$$

... a zase sucho

‘skeptik’ $P(T) = 0.01$

$$P(T|D_1) = \frac{1 \times 0.01}{1 \times 0.01 + 0.055 \times 0.99} = 0.155$$

$$P(T|D_2 \wedge D_1) = \frac{1 \times 0.155}{1 \times 0.155 + 0.055 \times 0.845} = 0.769$$

... po 4 deštích $P(T|D_4 \wedge D_3 \wedge D_2 \wedge D_1) = 0.984$

$$P(T|D_2 \wedge D_1) = \frac{P(D_2|T)P(T|D_1)}{P(D_2|T)P(T|D_1) + P(D_2|\bar{T})P(\bar{T}|D_1)}$$

$$P(T|D_2 \wedge D_1) = \frac{1 \times 0.948}{1 \times 0.948 + 0.055 \times 0.052} = 0.997$$

Podmíněná pravděpodobnost: příklad - obchodníci z deštěm

$$P(T|D_1) = \frac{P(D_1|T)P(T)}{P(D_1|T)P(T) + P(D_1|\bar{T})P(\bar{T})}$$

prší polovinu dní v roce

$$P(D_1|T) = 1$$

$$P(D_1|\bar{T}) = 0.5$$

$$P(T) = 0.5$$

$$P(T|D_1) = \frac{1 \times 0.5}{1 \times 0.5 + 0.5 \times 0.5} = 0.667$$

