Statistika pro informatiku

Souhrn látky

únor 2014

Obsah

1 Základy statistiky a pravděpodobnosti **2**
- 1.1 Rozlišení pojmů . . . 2
- 1.2 Pravděpodobnost jevu a jeho doplňku . . . 2
- 1.3 Sjednocení jevů . . . 2
- 1.4 Podmíněná pravděpodobnost . . . 3
- 1.5 Nezávislost jevů (průnik) . . . 3
- 1.6 Bayessova věta . . . 4

2 Rozdělení pravděpodobnosti **4**
- 2.1 Diskrétní rozdělení . . . 4
  - 2.1.1 Bernoulliho rozdělení . . . 4
  - 2.1.2 Binomické rozdělení . . . 4
  - 2.1.3 Geometrické rozdělení . . . 4
  - 2.1.4 Poissonovo rozdělení . . . 4
- 2.2 Spojité rozdělení . . . 4
  - 2.2.1 Rovnoměrné rozdělení . . . 4
  - 2.2.2 Exponenciální rozdělení . . . 4
  - 2.2.3 Normální (gaussovo) rozdělení . . . 5

3 Entropie **5**
- 3.1 Diskrétní veličiny . . . 5
- 3.2 Spojité veličiny . . . 5
- 3.3 Sdružená entropie . . . 6
- 3.4 Podmíněná entropie . . . 6

1 Základy statistiky a pravděpodobnosti

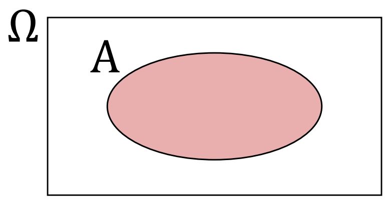
1.1 Rozlišení pojmů

Statistika TODO

Pravděpodobnost TODO

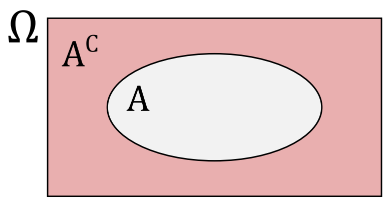
1.2 Pravděpodobnost jevu a jeho doplňku

$$P\left(A\right)=\frac{size\left(A\right)}{size\left(\Omega\right)}$$



Obrázek 1: Vennův diagram základní pravděpodobnosti jevu

$$P\left(A^C\right)=1-P\left(A\right)$$

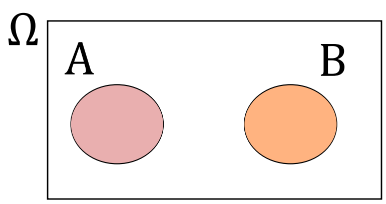


Obrázek 2: Vennův diagram doplňku jevu

1.3 Sjednocení jevů

Pro disjunktní jevy platí

$$P\left(A\cup B\right)=P\left(A\right)+P\left(B\right).$$

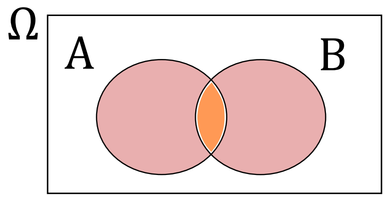


Obrázek 3: Disjunktní jevy

Pro jevy platní

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

Oblast průniku by byla započítána dvakrát, proto je potřeba ji odečíst.

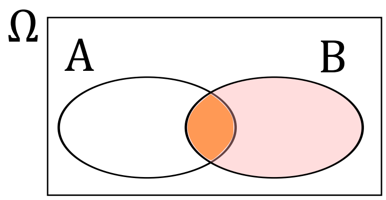


Obrázek 4: Jevy

1.4 Podmíněná pravděpodobnost

$$P(A|B) = \frac{P(A \cap B)}{P(B)},\ P(B) \neq 0$$

„Pravděpodobnost jevu A za podmínky, že jsme v B a že jev B nastal.“



Obrázek 5: Podmíněná pravděpodobnost

1.5 Nezávislost jevů (průnik)

Pro **nezávislé** jevy platí

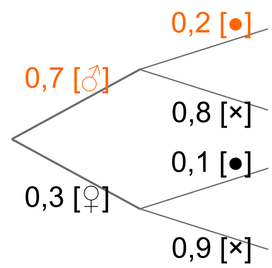
$$P(A \cap B) = P(A) * P(B).$$

Jinak platí

$$\begin{aligned} P(A \cap B) &= P(A|B)\, P(B) \\ P(A \cap B) &= P(B|A)\, P(A) \end{aligned}$$

1.6 Bayessova věta

$$P(♂|\bullet) = P(\bullet|♂) * P(♂) = P(♂) * P(\bullet|♂) = 0,7 * 0,2 = \underline{\underline{0,14}}$$



Obrázek 6: Bayessova věta pomocí stromu

2 Rozdělení pravděpodobnosti

2.1 Diskrétní rozdělení

2.1.1 Bernoulliho rozdělení

TODO

2.1.2 Binomické rozdělení

TODO

2.1.3 Geometrické rozdělení

$$\begin{array}{rcl} T & \sim & geom(p) \\ P(T > n) & = & (1-p)^n \\ P(T = k) & = & (1-p)^{k-1} * p \\ P(T \leq n) & = & 1 - (1-p)^n \end{array}$$

2.1.4 Poissonovo rozdělení

TODO

2.2 Spojité rozdělení

2.2.1 Rovnoměrné rozdělení

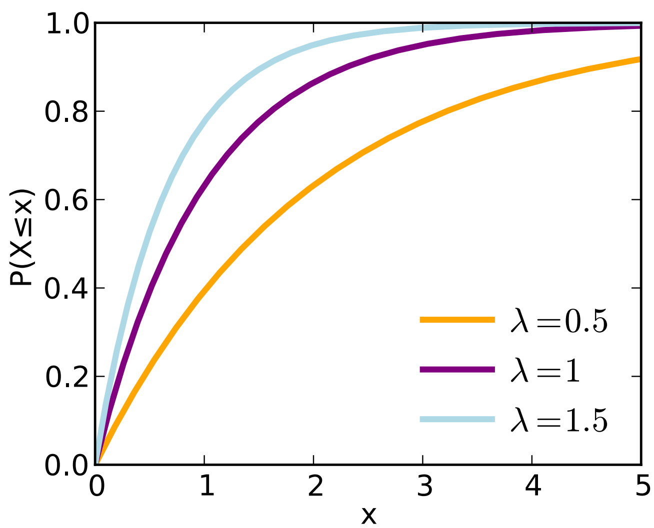
TODO

2.2.2 Exponenciální rozdělení

$$\begin{array}{rcl} X & \sim & exp(\lambda) \\ f(x) & = & \lambda e^{-\lambda x},\ x \geq 0 \\ P(X \leq x) & = & 1 - e^{-\lambda x} \end{array}$$

$$\begin{aligned} \mathrm{E}X &= \frac{1}{\lambda} \\ \mathrm{var}X &= \frac{1}{\lambda^2} \end{aligned}$$

$$P\left(\male|\bullet\right) = P\left(\bullet|\male\right) * P\left(\male\right) = P\left(\male\right) * P\left(\bullet|\male\right) = 0,7 * 0,2 = \underline{\underline{0,14}}$$



Obrázek 7: Graf distribuční funkce exponenciálního rozdělení[1]

2.2.3 Normální (gaussovo) rozdělení

TODO

3 Entropie

Pro entropii platí aditivita entropie

$$H\left(X,Y\right) = H\left(X\right) + H\left(Y|X\right).$$

3.1 Diskrétní veličiny

$$H\left(X\right) = -\sum p_i \log p_i$$

3.2 Spojité veličiny

$$H\left(X\right) = -\int f\left(x\right) \log f\left(x\right) \mathrm{d}x$$

3.3 Sdružená entropie

$$H(X,Y) = -\sum_{i,j} p_{i,j} \log p_{i,j}$$

3.4 Podmíněná entropie

$$H(X|Y) = -\sum_{i,j} p(x_i, y_i) \log p(y_i|x_i)$$

Reference

[1] The Free Encyclopedia Wikipedia. Exponential distribution. 2014.