# 3. cvičení

Pravděpodobnost a statistika 1, 1. 3. 2023 `https://iuuk.mff.cuni.cz/~chmel/2223/past/`

**Úloha 1** (Kos-ti-čky)
Máme tři normální šestistěnné hrací kostky a jednu speciální šestistěnnou kostku, kde jsou tři jedničky a tři dvojky. Vybereme uniformně náhodně jednu z kostek, hodíme a padne jednička. Jaká je pravděpodobnost, že jsme vybrali normální kostku?

**Řešení**
Jevy: $N$ - výběr normální kostky, $J$ - padla jednička. $P(N|J) = \frac{P(J|N)\cdot P(N)}{P(J)} = \frac{1/6\cdot 3/4}{\frac{1}{6}\cdot\frac{3}{4}+\frac{1}{2}\cdot\frac{1}{4}} = \frac{1/8}{1/4} = \frac{1}{2}$.

**Úloha 2** (Spam or ham?)
Petr dostává hodně emailů, ale 80 % z nich jsou spamy. Jeho spamový filtr 90 % spamů správně označí, ale také 5 % řádných emailů označí jako spam.

1. Kolik procent emailů bude označeno jako spamy?
2. Kolik procent řádných emailů je mezi těmi, co jsou označené jako spamy?
3. Kolik procent spamů je mezi emaily, které testem prošly?

**Řešení**
Značíme jevy: $O$: „označený jako spam“, $S$ : „je spam“.

1. $P(O) = P(O|S)\cdot P(S) + P(O|\neg S)\cdot P(\neg S) = 0.9\cdot 0.8 + 0.05\cdot 0.2 = 0.73$
2. $P(\neg S|O) = \frac{P(O|\neg S)\cdot P(\neg S)}{P(O)} = \frac{0.05\cdot 0.2}{0.73} = \frac{1}{73}$
3. $P(S|\neg O) = \frac{P(\neg O|S)\cdot P(S)}{P(\neg O)} = \frac{0.1\cdot 0.8}{0.27} = \frac{8}{27}$

## Nezávislé jevy

**Úloha 3** (Nezávislost a doplňky)
Ukažte, že jsou-li jevy $A$, $B$ nezávislé, pak jsou nezávislé i $A$, $B^c$ a také $A^c$, $B^c$.

**Řešení**
$P(A\cap B^c) = P(A) - P(A\cap B) = P(A) - P(A)\cdot P(B) = P(A)\cdot(1-P(B)) = P(A)\cdot P(B^c)$.
Dva doplňky analogicky pro $A = B^c, B = A$.

**Úloha 4** (Nezávislost a disjunktnost)
Mohou být dva jevy nezávislé a zároveň disjunktní?

**Řešení**
Pokud $A, B$ jsou disjunktní, pak $P(A\cap B) = 0$, a tedy pro nezávislost nutně musí platit $P(A)P(B) = 0$, což nastane jen tehdy, když $P(A) = 0 \vee P(B) = 0$. Tedy ano, můžou, ale jen tehdy, když alespoň jeden z těchto jevů je nemožný (má nulovou pravděpodobnost).

**Úloha 5** (Součin pravděpodobností a nezávislost po dvou)
Najděte jevy $A$, $B$, $C$ takové, že $P(A\cap B\cap C) = P(A)P(B)P(C)$, ale jevy nejsou po dvou nezávislé.

**Řešení**
Liná verze: házeme šestistěnnou kostkou, jevy jsou: $A$: „hodíme sedmičku“ (formálně $A = \emptyset$), $B = \{1,3,5\}$: „hodíme liché číslo“, $C = \{1,2,3\}$: „hodíme číslo $\leq 3$“. $P(A\cap B\cap C) = P(A\cap B) = P(A\cap C) = P(A)\cdot P(B) = P(A)\cdot P(C) = P(A)\cdot P(B)\cdot P(C) = 0, P(B\cap C) = 1/3 \neq \frac{1}{4} = P(B)\cdot P(C)$.

## Rychlý pohled na náhodné veličiny

**Úloha 6** (Pokusy)
Prokop hází basketbalovým míčem na koš, v každém pokusu má pravděpodobnost zásahu $1/10$, pokusy jsou nezávislé. Skončí po první trefě. Označme $X$ celkový počet hodů.

1. Jaká je $P[X > k]$?
2. Jaká je distribuce $X$? Tj. určete pravděpodobnostní funkci $p_X$, tj. pro každé $x$ určete $P[X = x]$.
3. Jaká je $P[X \geq 10 | X \geq 5]$?

**Řešení**

1. $P[X > k] = (1 - 1/10)^k$
2. $P[X = k] = (1 - \frac{1}{10})^{k-1} \frac{1}{10}$
3. $P[X \geq 10 | X \geq 5] = \frac{(1-1/10)^9}{(1-1/10)^4} = (1 - 1/10)^5$

---

**Bonusové úlohy**

**Úloha 7** (Prosecutor's fallacy)
Paní C. umřely dvě děti krátce po narození. Je obžalovaná za dvojnásobnou vraždu. Žalobce argumentuje takto: Pravděpodobnost syndromu náhlého úmrtí kojenců je $1/8500$. Takže pravděpodobnost dvou takových jevů je $1/8500^2$. Tudíž pravděpodobnost, že paní C. je nevinná, je $1/8500^2$, což je hodně málo. Formulujte argumenty žalobce v řeči pravděpodobnosti a nalezněte v nich dvě chyby.

**Řešení**
Pravděpodobnosti nemusí být nezávislé, a tohle je pst dvou náhlých úmrtí za předpokladu, že je nevinná. Naopak, jaká je pst toho, že je nevinná, za předpokladu dvou úmrtí je těžké říct.

## Na procvičení

**Úloha 8** (Kouřové signály)
Kouřovými signály přenášíme binární soubor. Je proto poměrně vysoká pravděpodobnost chyby u každého bitu: 0 se jako 0 přenese jen s pravděpodobností 0.9, 1 jako 1 jen s pravděpodobností 0.8. Předpokládejme (trochu neseriózně), že jednotlivé znaky se přenáší nezávisle. Dále předpokládejme, že ve vysílané zprávě je stejně nul a jedniček.

1. Pokud jsme dostali signál 0, jaká je pravděpodobnost, že byl opravdu vyslán?
2. Dostali jsme zprávu 0010. Jaká je pravděpodobnost, že byla opravdu vyslána?
3. Jak se výpočet změní, pokud budeme pro kontrolu vysílat každý symbol třikrát (a pak vezmeme častější z těch tří pokusů)?

Úlohu si můžete zjednodušit předpokladem, že 0 a 1 mají stejnou pravděpodobnost správného přenosu.

**Řešení**
TODO

**Úloha 9** (Nemocnice na kraji města)
Na chorobu $C$ máme dva testy, $A$ a $B$. Test $A$ má senzitivitu i specificitu $p = 0.95$ (senzitivita je pravděpodobnost, že nemocného člověka test opravdu určí jako nemocného, tedy $P$[test řekne „nemocný“|člověk je nemocný] a specificita je $P$[test řekne „zdravý“|člověk je zdravý], tedy pravděpodobnost, že zdravého člověka test opravdu označí za zdravého). Test $B$ vždy řekne, že pacient je zdravý. Předpokládejte, že $P(C) = 0.01$.

1. Spočtěte pro oba testy pravděpodobnost úspěchu (tj. správné odpovědi), použijeme-li je na náhodného pacienta. Co to říká o užitečnosti obou testů?
2. Pro jaké $p$ je pravděpodobnost úspěchu obou testů stejná?

**Řešení**

Pro $A: P(S) = P(C|A)\cdot P(A) + P(C^c|A^c)\cdot P(A^c) = \frac{P(A|C)\cdot P(C)}{P(A)}\cdot P(A) + \frac{P(A^c|C^c)\cdot P(C^c)}{P(A^c)}\cdot P(A^c) = P(A|C)\cdot P(C) + P(A^c|C^c)\cdot P(C^c) = p\cdot 0.01 + p\cdot 0.99 = p$, pro $B$ analogicky $P(B|C)\cdot P(C) + P(B^c|C^c)\cdot P(C^c) = 0 + 0.99 = 0.99$. Stejná pravděpodobnost je pro $p = 0.99$.

**Úloha 10** (Exit poll)

Ve volbách hlasují lidé pro dva kandidáty, $A$ a $B$. Při odchodu z volební místnosti jsou voliči náhodně požádáni o účast v exit-poll. Předpokládejme, že kdo odpoví, odpoví popravdě koho volil, ale ne všichni se zúčastní. Označíme-li $E$ množinu voličů, kteří se exit-pollu zúčastní, tak předpokládejme $P(E|A) = 0.7$ a $P(E|A^c) = 0.4$. Výsledky exit-pollu jsou 60 % pro $A$. Jaký je skutečný podíl lidí, kteří hlasovali pro $A$?

**Řešení**

Tohle tedy znamená, že $P(A|E) = 0.6$, a nás by zajímalo zjistit $P(A), P(B)$ – budeme předpokládat $B = A^c$ pro názornost, dále označíme $a = P(A)$ pro zkrácení, navíc s $P(A)$ pracujeme jako s neznámou. Pak tedy $a = P(A) = P(A|E)\cdot P(E) = P(A|E)\cdot(P(E|A)\cdot P(A) + P(E|B)\cdot P(B)) = P(A|E)P(E|A)a + P(A|E)P(E|B)(1-a) \rightsquigarrow a = 0.42a + 0.24(1-a) \rightsquigarrow P(A) = 12/41$.

**Úloha 11** (Svědek)

Jirka je policista v New Vegas a vyšetřuje loupež. Lupič po krádeži skočil do taxíků a ujel. Svědek na místě Jirkovi řekl, že skočil do žlutého taxíku.

Jirka je ale zároveň obeznámený s Bayesovou větou, a tak by ho zajímalo, jak moc může výpovědi věřit. Konkrétně ví, že v New Vegas je 80 % taxíků černých a zbylých 20 % je žlutých. Zároveň ví (z kurzů kriminalistiky), že svědek dobře identifikuje a zapamatuje si barvu taxíku s pravděpodobností 80 %. Jaká je pravděpodobnost toho, že svědek vypověděl pravdivě?

**Řešení**

Dva jevy: $T$ - taxík byl žlutý, $V$ - vypověděl žlutou. $P(T|V) = \frac{P(V|T)\cdot P(T)}{P(V)} = \frac{0.8\cdot 0.2}{0.8\cdot 0.2 + 0.2\cdot 0.8} = 1/2$.

**Úloha 12** (Po dvou nezávislé, ale ne nezávislé)

Najděte jevy $A$, $B$, $C$ takové, že jsou po dvou nezávislé, ale $P(A\cap B\cap C) \neq P(A)P(B)P(C)$.

**Řešení**

Prostor: dvě mince. Jevy: $A$: na první minci orel, $B$: na druhé minci orel, $C$: na obou mincích dohromady jeden orel.

Snadno se spočítá, že $P(A) = P(B) = P(C) = 1/2, P(A\cap B) = P(A\cap C) = P(B\cap C) = 1/4$, $P(A\cap B\cap C) = 0$.

---

# Domácí úkol 3

`https://iuuk.mff.cuni.cz/~chmel/2223/past/` Zadáno 1. 3. 2023

Pravděpodobnost a statistika 1 Odevzdejte do 8. 3. 2023 9:00 přes Poštovní sovu

---

**Úloha 1** (Truhla)

V truhle je sto mincí. Z nich je 99 normálních, ale jedna má na obou stranách orla. Řekneme našemu kamarádovi, aby pomocí Uniformně Náhodného Vytahovače™ jednu minci z truhly vytáhl a šestkrát s ní hodil. Zjistili jsme, že nám šestkrát padnul orel. Jaká je pravděpodobnost, že kamarád vytáhnul minci se dvěma orly?