Klasická pravděpodobnost

Klasická definice pravděpodobnosti:

- Ω je množina všech možných výsledků náhodného pokusu
- $\omega \in \Omega$ elementární jev
- $A \subset \Omega$ náhodný jev
- Nechť Ω obsahuje konečný počet prvků tj. $\Omega = \{\omega_1, \dots, \omega_n\}$ a nechť ω_i jsou stejně pravděpodobné. Pak pravděpodobnost náhodného jevu A definujeme jako $P(A) = \frac{|A|}{|\Omega|} = \frac{|A|}{n}$, kde $|A| =$ počet prvků množiny A.

Vlastnosti:

- $0 \leq P(A) \leq 1$,
- $P(A^c) = 1 - P(A)$,
- jestliže $A \subset B$, pak $P(A) \leq P(B)$ a $P(B \setminus A) = P(B) - P(A)$,
- $P(A \cap B) = P(A) - P(A \cap B^c)$,
- $P(A \cup B) = P(A) + P(B) - P(A \cap B)$,
- princip inkluze a exkluze

$$P(A_1 \cup A_2 \cup \dots \cup A_n) = \sum_{i=1}^{n} P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \dots + (-1)^{n+1} P(A_1 \cap A_2 \cap \dots \cap A_n).$$

1. Házíme čtyřmi šestistěnnými hracími kostkami. Určete, jaká je pravděpodobnost, že

   a) padnou čtyři různá čísla,
   b) padnou pouze lichá čísla,
   c) součet čísel na všech kostkách dohromady bude roven 6,
   d) součet čísel bude větší než 5,
   e) padne alespoň jedna šestka.

2. V regálu je 6 lahví normálního rumu a 4 lahve pančovaného rumu (vizuálně k nerozeznání). Náhodně vybereme z regálu 3 lahve a z každé ochutnáme. Určete, s jakou pravděpodobností

   a) byl právě ve dvou námi ochutnaných lahvích methanol,
   b) byl alespoň v jedné námi ochutnané lahvi methanol.

3. Na svazku máme 8 různých klíčů a pokoušíme se odemknout zámek. Vyzkoušený klíč vždy dáme stranou a náhodně vybereme další klíč ze zbývajících. Jaká je pravděpodobnost, že odemkneme až na pátý pokus?

4. Uvažujme n různých dopisů a n různých obálek (s již nadepsanou adresou). Zmatená sekretářka umístí dopisy do obálek zcela náhodně.

   a) Jaká je pravděpodobnost, že je alespoň jeden dopis ve správné obálce?
   b) Spočtěte limitu této pravděpodobnosti pro $n \to \infty$.

5. (Maxwell-Boltzmannovo schéma) Do vlaku s n vagóny nastupuje r cestujících. Předpokládejme, že každý člověk si vybírá vagón zcela náhodně.

   a) Určete, s jakou pravděpodobností bude v prvním vagóně právě $k \leq r$ cestujících.
   b) Jaká je pravděpodobnost, že žádný vagón nebude prázdný?
   c) Spočítejte limitu pravděpodobnosti z bodu (a) pro $n \to \infty, r \to \infty$ tak, že $r/n \to \lambda > 0$.

6. (Bose-Einsteinovo schéma) Babička rozděluje r tisícikorun do n obálek pro svých n vnoučat k Vánocům. Peníze rozmístí náhodně (všechna rozmístění jsou stejně pravděpodobná).

   a) Určete pravděpodobnost, že vnuk Karel dostane právě k tisícikorun.
   b) Jaká je pravděpodobnost, že každé z vnoučat dostane alespoň nějaké peníze?
   c) Spočtěte limitu pravděpodobnosti z bodu $n \to \infty$, $r \to \infty$ tak, že $r/n \to \lambda > 0$.

Geometrická pravděpodobnost

1. Na úsečce délky l jsou náhodně umístěny dva body, které tuto úsečku rozdělí na tři části. S jakou pravděpodobností je možné z takto vzniklých tří úseček sestrojit trojúhelník?

2. (1777 - Buffonova jehla - Georges Louis Leclerc de Buffon) Na podlaze jsou parkety šířky l. Na podlahu spadne jehla délky $r < l$. Jaká je pravděpodobnost, že bude jehla ležet na dvou parketách? Šířku mezery mezi parketami zanedbáváme.

3. (Bertrandův paradox - Joseph Louis François Bertrand) Uvažujme kružnici a zvolme náhodně tětivu této kružnice. Jaká je pravděpodobnost, že délka této tětivy bude větší než délka strany rovnostranného trojuhelníka vepsaného do této kružnice?

Podmíněná pravděpodobnost, nezávislost jevů

Nechť A, B jsou náhodné jevy, $P(B) > 0$. **Podmíněnou pravděpodobnost** jevu A za podmínky B definujeme jako

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

Nezávislost. Náhodné jevy A, B jsou navzájem nezávislé, jestliže platí

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

Náhodné jevy $A_1, \dots A_n$ jsou nezávislé, jestliže pro každé $r \leq n$ a každou $\{i_1, \dots, i_r\}$ podmnožinu $\{1, \dots, n\}$ platí

$$P(A_{i_1} \cap \dots \cap A_{i_r}) = P(A_{i_1}) \cdot \dots P(A_{i_r}).$$

(Tj. součinovou podmínku musíme ověřit pro všechny dvojice, všechny trojice ... atd.)

Věta o úplné pravděpodobnosti:
Nechť $A, B_1, B_2, \dots$ jsou náhodné jevy takové, že $B_i \cap B_j = \emptyset$ pro všechna $i \neq j$, $\bigcup_i B_i = \Omega$ a $P(B_i) > 0$ pro všechna $i = 1, 2, \dots$. Pak

$$P(A) = \sum_i P(A \cap B_i) = \sum_i P(A|B_i)P(B_i).$$

Bayesova věta:
Nechť $A, B_1, B_2, \dots$ jsou náhodné jevy takové, že $B_i \cap B_j = \emptyset$ pro všechna $i \neq j$, $\bigcup_i B_i = \Omega$, $P(B_i) > 0$ pro všechna i, a nechť $P(A) > 0$. Pak

$$P(B_i|A) = \frac{P(B_i \cap A)}{P(A)} = \frac{P(A|B_i)P(B_i)}{\sum_j P(A|B_j)P(B_j)}.$$

Věta o násobení pravděpodobností:
Jestliže náhodné jevy $A_1, \dots, A_n$ splňují $P(\cap_{i=1}^n A_i) > 0$, pak

$$P(\cap_{i=1}^n A_i) = P(A_n | \cap_{i=1}^{n-1} A_i) \cdot P(A_{n-1} | \cap_{i=1}^{n-2} A_i) \cdot \dots \cdot P(A_2|A_1)P(A_1).$$

Příklady 9. a 10. jsou včetně řešení převzaty z: http://cmp.felk.cvut.cz/∼navara/pms/PMScvic.pdf

1. Házíme dvěma pravidelnými kostkami.
   a) Jaká je pravděpodobnost, že padla šestka, za podmínky, že celkový součet je 8?
   b) Jsou jevy [padla šestka] a [celkový součet je 8] nezávislé?
2. Házíme dvěma pravidelnými kostkami - modrou a zelenou. Označme jevy A =[na modré kostce padlo sudé číslo], B =[na zelené kostce padlo liché číslo], C =[součet čísel je lichý].
   a) Jsou náhodné jevy A,B,C po dvou nezávislé?
   b) Jsou jevy A,B,C nezávislé?
3. Házíme dvěma pravidelnými kostkami najednou, dokud nepadne součet 5 nebo součet 7 (na obou kostkách dohromady). S jakou pravděpodobností padne dříve součet 5 než součet 7?
4. Dne 9.10.2020 byla ve večerních zprávách zveřejněna následující zpráva: Ministr zdravotnictví Roman Prymula zvažuje možnost otestování celého národa na COVID 19. Byla uvedena následující data:
   - počet testovaných jedinců . . . 10 000 000,
   - falešně negativních jedinců . . . 20 000,
   - pozitivních jedinců . . . 180 000,
   - falešně pozitivních jedinců . . . 249 000.

   a) Pokud by testování dopadlo dle odhadu a konkrétní pacient by byl testem označen jako pozitivní, jaká je pravděpodobnost, že je skutečně pozitivní?
   b) Určete pravděpodobnost, že bude zdravý jedinec označen testem jako pozitivní.
5. Ve třídě je 70% chlapců a 30% dívek. Dlouhé vlasy má 10% chlapců a 80% dívek.
   a) Jaká je pravděpodobnost, že má náhodně vybraná osoba dlouhé vlasy?
   b) Vybraná osoba má dlouhé vlasy. Jaká je pravděpodobnost, že je to dívka?
6. Roztržitý profesor zapomene v obchodě deštník s pravděpodobností 1/4 (za předpokladu, že s deštníkem do obchodu přišel). Cestou z Karlína navštívil čtyři obchody a domů přišel bez deštníku. Jaká je pravděpodobnost, že deštník zapomněl ve čtvrtém obchodě?
7. V jednom ze tří trezorů je skryta odměna, ostatní dva jsou prázdné. Soutěžící je vyzván, aby si vybral trezor, ve kterém myslí, že je skryta odměna. V případě, že se trefí, tak odměnu získá. Po vybrání je otevřen jeden ze dvou dalších trezorů, ale vždy pouze ten, který je prázdný, o čemž je soutěžící informován. Poté je soutěžící vyzván, zda chce změnit volbu trezoru a vybrat si druhý zbývající. Je pro soutěžícího změna trezoru výhodná?
8. Skříňka má 3 zásuvky. V první jsou 2 zlaté mince, ve druhé jedna zlatá a jedna stříbrná, v třetí 2 tříbrné mince. Zvolíme náhodně jednu zásuvku, z ní vytáhneme naslepo jednu minci. Jaká je pravděpodobnost, že v zásuvce zbyde zlatá mince, jestliže vytažená mince byla stříbrná?
9. Požití alkoholu bylo prokázáno u 1% všech řidičů a u 10% řidičů, kteří způsobili dopravní nehodu. Kolikrát se požitím alkoholu zvyšuje riziko nehody?
10. Z 60 žijících členů klubu vysloužilých námořních kapitánů jich 5 zažilo ztroskotání (jednou). Podle statistiky při ztroskotání lodi v této oblasti třetina kapitánů zahyne. Odhadněte pravděpodobnost, že kapitán zažije ztroskotání (aspoň jednou za život – možnost opakovaného ztroskotání téhož kapitána i předčasného úmrtí z jiné příčiny zanedbáváme).
11. Náhodně vybereme kladné celé číslo N, rozdělení pravděpodobnosti je $P[N = i] = 2^{-i}$. Poté hodíme N kostkami. Nechť S je součet hodnot, které při hodu padly. Určete podmíněné pravděpodobnosti $P[N = 2|S = 4]$ a $P[S = 4|N \text{ je sudé}]$.
12. a) Rodina má dvě děti, starší je dcera. Jaká je pravděpodobnost, že mají dvě dcery?
    b) Rodina má dvě děti, (aspoň) jedno z nich je dcera. Jaká je pravděpodobnost, že mají dvě dcery?
13. Jazykový korektor změní 99% chybných slov na správná a 0.01% správných na chybná. Změnil 2% slov. Odhadněte množství chybných slov v jeho výstupu.

Diskrétní rozdělení

Alternativní rozdělení $X \sim Alt(p)$**:** $P(X=1)=p, \quad P(X=0)=(1-p)$, $\mathbb{E}X=p$, $\text{var}X=p(1-p)$.

Použití: Sledujeme jednu realizaci pokusu (s pravděpodobností úspěchu p), $X=1$ v případě úspěšného pokusu a $X=0$ v opačném případě.

Geometrické rozdělení $X \sim Geom(p)$**:** $P(X=k)=p(1-p)^k,\ k=0,1,2,...$, $\mathbb{E}X=\frac{1-p}{p}$, $\text{var}X=\frac{1-p}{p^2}$.

Použití: Sledujeme počet neúspěšných pokusů před prvním úspěšným při neustálém opakování stejného pokusu (který má pravděpodobnost úspěchu p). Předpokládáme, že výsledky jednotlivých pokusů jsou na sobě nezávislé.

Binomické rozdělení $X \sim Bi(n,p)$**:** $P(X=k)=\binom{n}{k}p^k(1-p)^{n-k},\ k=0,1,2,...,n$, $\mathbb{E}X=np$, $\text{var}X=np(1-p)$.

Použití: Uvažujeme n nezávislých náhodných pokusů (každý skončí úspěchem s pravděpodobností p). Sledujeme, kolik bude úspešných pokusů z těchto n celkových pokusů.

Poissonova rozdělení $X \sim Po(\lambda)$**:** $P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!},\ k=0,1,2,...$, $\mathbb{E}X=\lambda$, $\text{var}X=\lambda$.

Použití: Uvažujeme události, které přicházejí nezávisle na sobě (například hovory v telefonní ústředně, či příchody jednotlivých zákazníků do obchodu). Počet těchto událostí (většinou uvažujeme situaci, kdy nás zajímají události, které nastanou v rámci nějakého časového intervalu) pak má Poissonovo rozdělení. Parametr λ většinou určujeme z průměrného počtu událostí, které v sledovaném časovém intervalu nastaly (λ je roven tomuto průměrnému počtu).

Spojitá rozdělení

Rovnoměrné rozdělení $X \sim R(a,b)$**:** $f(x)=\frac{1}{b-a},\ x\in(a,b)$, $\mathbb{E}X=\frac{a+b}{2}$, $\text{var}X=\frac{(b-a)^2}{12}$.

Použití: Používá se pro situace, kdy náhodný výsledek může být libovolná hodnota z intervalu (a,b) a nemáme důvod jakoukoli z těchto hodnot preferovat. Například čekací doba na spoj, který má pravidelné intervaly.

Exponenciální rozdělení $X \sim Exp(\lambda)$**:** $f(x)=\lambda e^{-\lambda x},\ x>0$, $\mathbb{E}X=\frac{1}{\lambda}$, $\text{var}X=\frac{1}{\lambda^2}$.

Použití: Pro časy mezi událostmi, které přicházejí nezávisle na sobě a jejich střední počet je přímo úměrný délce sledovaného intervalu. Například čas mezi dvěma telefonáty v telefonní ústředně. Jde o rozdělení bez paměti, tj. platí $P(X>t+s|X>s)=P(X>t)$.

Normální rozdělení $X \sim N(\mu,\sigma^2)$**:** $f(x)=\frac{1}{\sqrt{2\pi\sigma^2}}e^{-\frac{(x-\mu)^2}{2\sigma^2}},\ x\in\mathbb{R}$, $\mathbb{E}X=\mu$, $\text{var}X=\sigma^2$.

Použití: Pro náhodnou veličinu, která je součtem mnoha nezávislých náhodných veličin. Pro hodnotu, která je výslednicí velkého množství náhodných vlivů. Například výška dospělého muže, nebo inteligenční kvocient náhodně vybrané osoby.

1. Nechť $X \sim Po(\lambda)$. Odvoďte vzorec pro střední hodnotu $\mathbb{E}X$ a rozptyl var$X$ bez použití znalostí o střední hodnotě a rozptylu binomického rozdělení.

2. Uvažujme náhodnou veličinu X s hustotou $f_X(x) = \lambda e^{-\lambda x}, \quad x > 0$.

   a) Určete střední hodnotu $\mathbb{E}X$.

   b) Určete rozptyl var$X$.

   c) Určete pravděpodobnost, že $X > t, \quad t > 0$.

   d) Jaká je pravděpodobnost, že $X < t + s$ za podmínky $X > t$ $(t > 0, s > 0)$? Jinak řečeno, je-li například X doba čekání na příjezd autobusu, tak jaká je pravděpodobnost, že budeme celkem čekat mémě než $s + t$ minut, za podmínky, že jsme již čekali t minut?

3. Univerzitní hokejový tým UK Praha vyšle na branku soupeře průměrně 42 střel na utkání. Průměrně každá osmá střela končí gólem.

   a) Určete pravděpodobnost, že padne první gól zmíněného týmu až při desátém střeleckém pokusu.

   b) Pokud vystřelil tým UK Praha v první třetině patnáctkrát, jaká je pravděpodobnost, že dal právě dva góly?

   c) Jaká je pravděpodobnost, že vystřelí zmíněný tým v první třetině právě patnáct střel?

   d) Jaká je pravděpodobnost, že budeme čekat na první střelu týmu UK více jak tři minuty?

4. Cyril má na svazku 8 klíčů a snaží se odemknout dveře (ke kterým pasuje právě jeden klíč). Náhodně vybere klíč a vyzkouší ho. Po každém neúspěšném pokusu mu klíče spadnou na zem a další klíč znovu volí zcela náhodně. Tak pokračuje, dokud konečně dveře neotevře.

   a) Jaké je rozdělení počtu všech neúspěšných Cyrilových pokusů?

   b) Jaký je stědní počet neúspěšných pokusů?

   c) Jaký je rozptyl počtu neúspěšných pokusů?

5. Biatlonový závodník během závodu střílí dvacetkrát na terč. Pravděpodobnost, že při konkrétní střele mine, je 0.09. Předpokládáme, že výsledky jednotlivých pokusů jsou nezávislé a za každý minutý terč dostává závodník trestnou minutu.

   a) Jaká je pravděpodobnost, že závodník dostane během závodu více jak dvě trestné minuty?

   b) Jaký je očekávaný počet trestných minut, který závodník získá?

   c) Jaký počet získaných trestných minut je nejpravděpodobnější?

6. Nechť počet příchozích hovorů na ústřednu během daného časového intervalu se řídí Poissonovým rozdělením (parametr λ je přímo úměrný délce časového intervalu). Průměrně přijde během hodiny dvacet hovorů.

   a) Jaké je rozdělení počtu příchozích hovorů během deseti minut?

   b) Jaká je pravděpodobnost, že během následující minuty přijme ústředna alespoň dva hovory?

7. Semena mají klíčivost $p \in (0, 1)$. Jaký je optimální počet n semen v jamce, aby byla co nejvyšší pravděpodobnost, že vyklíčí právě jedno? Řešte obecně a pro $p = 1/3$.

8. Do obchodu přijde průměrně 10 zákazníků za hodinu, z toho je průměrně 60 % žen. Jaká je pravděpodobnost, že během půl hodiny přijdou do obchodu alespoň 3 zákazníci a všechno to budou ženy?

9. Nechť se počet příchozích hovorů na ústředně během daného časového intervalu řídí Poissonovým rozdělením (parametr λ je přímo úměrný délce časového intervalu). Průměrně přijde během hodiny dvacet hovorů.

   a) Jaké je rozdělení doby čekání na další hovor (v minutách)? Určete jeho hustotu.

   b) Jaké je rozdělení doby čekání na další hovor, víme-li, že hovor nepřišel během časového intervalu $[0, T]$?

   c) Na jak dlouho si operátor v ústředně může odskočit pro kafe, aby s pravděpodobností 0.9 nepromeškal žádný příchozý hovor?

10. Mějme náhodnou veličinu X s rozdělením s hustotou $f(x) = Ce^{-\frac{(x-\mu)^2}{2\sigma^2}}$ pro $x \in \mathbb{R}$.

a) Určete konstantu C.
b) Odvoďte vzorec pro střední hodnotu tohoto rozdělení.
c) Odvoďte vzorec pro rozptyl tohoto rozdělení.

11. Nechť $X \sim N(1.5, 6)$.S využitím znalosti distrubuční funkce normovaného normálního rozdělení $\Psi(x)$.

    a) $P(X < 2.5)$.
    b) $P(X < 1)$.
    c) $P(1.5 < X < 4)$.

12. Nechť $X \sim N(0, 1)$. Určete $\mathbb{E}X^k$ pro $k = 1, 2, ...$

Nezávislost náhodných veličin

Nezávislost náhodných veličin. Náhodné veličiny $X_1, X_2, ..., X_n$ jsou **nezávislé**, jestliže pro všechna $x_1, ..., x_n \in \mathbb{R}$ platí rovnost

$$P(X_1 < x_1, X_2 < x_2, ..., X_n < x_n) = P(X_1 < x_1) \cdot P(X_2 < x_2) \cdot ... \cdot P(X_n < x_n).$$

Věta

- Diskrétní náhodné veličiny $X_1, X_2, ..., X_n$ jsou nezávislé právě tehdy, když

  $$P(X_1 = x_1^{(i)}, \ldots, X_n = x_n^{(i)}) = \Pi_{j=1}^{n} P(X_j = x_j^{(i)})$$

  pro všechna $x_1, ..., x_n \in \mathbb{R}$.

- Spojité náhodné veličiny $X_1, X_2, ..., X_n$ jsou nezávislé právě tehdy, když

  $$f_{\mathbb{X}}(x_1, x_2 \ldots, x_n) = f_{X_1}(x_1) \cdot f_{X_2}(x_2) \ldots f_{X_n}(x_n),$$

  pro skoro všechna $\mathbf{x} = (x_1, x_2 \ldots, x_n) \in \mathbb{R}^n$

1. Nechť X_1 má Poissonovo rozdělení s parametrem λ_1, X_2 má Poissonovo rozdělení s parametrem λ_2 a X_1 a X_2 jsou nezávislé náhodné veličiny. Určete rozdělení náhodné veličiny $Y = X_1 + X_2$.

2. Na autě jsou prováděny dvě nezávislé opravy a obě opravy budou hotovy do jedné hodiny. Předpokládejme, že obě opravy jsou v takové fázi, že rozdělení času do ukončení konkrétní opravy je rovnoměrné.

   a) Jaká je střední hodnota a rozptyl čekání na ukončení první opravy?

   b) Jaké je rozdělení doby do ukončení obou oprav?

   c) Určete pravděpodobnost, že obě opravy budou ukončeny do 45 minut.

   d) Uvažujme situaci, kdy opravy nebudou prováděny současně, ale postupně. Jaké rozdělení bude mít čas ukončení obou oprav?

   e) Uvažujme nyní n nezávislých oprav, jejichž doba má rovnoměrné rozdělení $R(0, 60)$. Nechť T_1 je čas, kdy bude dokončena první z oprav, a T_2 čas, kdy bude dokončena poslední z oprav. Jaké je rozdělení nahodných veličin T_1 a T_2?

3. Nechť má náhodný vektor (X, Y) rovnoměrné rozdělení na množině $\Omega = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$.

   a) Určete sdruženou hustotu rozdělení vektoru (X, Y).

   b) Určete marginální hustotu $f_X(x)$ rozdělení nahodne veličiny X.

   c) Jsou náhodné veličiny X a Y nezávislé?

4. Nechť má náhodný vektor (X, Y) rozdělení s hustotou
   $f_{X,Y}(x, y) = \frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{x^2+2\rho xy+y^2}{2(1-\rho^2)}}$ na $\mathbb{R}^2$.

   a) Určete rozdělení nahodne veličiny X.

   b) Jsou nahodne veličiny X a Y nezávislé?

5. Oštěpařka Anna má průměrnou délku hodu 67 m a směrodatnou odchylku 6 m. Oštěpařka Barbora má průměrnou délku hodu 75 m, a a směrodatnou odchylku 3 m. Předpokládejme, že délky hodů mají nezávislá normalní rozdělení. Spočtěte pravděpodobnost, že při jednom hodu hodí Anna dál než Barbora.

6. Uvažujme nezávislé náhodné veličiny X_i, kde $X_i \sim R(0, a)$.

   a) Nechť $Y = \frac{2}{n} \sum_{i=1}^{n} X_i$. Určete $\mathbb{E}Y$ a var$Y$.

   b) Nechť $Z = \max\{X_1, ..., X_n\}$. Určete $\mathbb{E}Z$ a var$Z$.

   c) Zamyslete se nad výsledky z částí a) a b). Co lze odhadnout o konvergnenci náhodných veličin Y a Z pro $n \to \infty$?