3 Opakované pokusy

3.1 Základy psti

- každý experiment je založen na *náhodném pokusu*
  - výzkum: výška člověka: náhodný pokus ... změříme 1 člověka;
  - výzkum: oblíbená značka auta ... zeptáme se jednoho náhodného muže;
  - výzkum: které číslo padne na kostce ... hodíme kostkou;
- *základní prostor* Ω ... množina všech možných výsledků
  - výška člověka ... $0 - \infty$; $0 - 4m$;
  - auta ... škoda, bmw, mercedes, mazda, VW, hyundai, jiné;
  - kostka ... 1-6;
- *možné výsledky* ω ... prvky základního prostoru
  - náhodný pokus - hodím kostkou - základní prostor = 6 možností : $\Omega = \{1, 2, 3, 4, 5, 6\}$ - možné výsledky: $\{\omega_1 = 1; \omega_2 = 2; \ldots \omega_6 = 6\}$
- hodila jsem kostkou: nastal *jev*:
  - padla 1, padla 2, ... padla 6
  - padlo liché číslo
  - *jev jistý:* padne buď 1,2,3,4,5,nebo 6
  - *jev opačný:* K jevu A:(padne 1-3) je opačný jev A':(padne 4-6)
  - *jevy neslučitelné:* padne 1 a padne sudé číslo
- *pravděpodobnost (pst)* = jak velká je naděje, že nějaký jev nastane
  - $\Pr(A) = \Pr(nastal jev A)$
  - $\Pr(A) \in \langle 0; 1\rangle$; resp. $\langle 0\,\% - 100\,\%\rangle$
  - příklad: hodím kostkou:
    * Pr(padne 1) = 1/6; 16.7 %
    * Pr(liché číslo) = 1/2; 50 %
    * Pr(padne 1,2,3,4,5 nebo 6) = 1; 100 %
  - Vlastnosti psti:
    * $\Pr(A) \geq 0$ (nezáporná).
    * Pr(jistý jev) = 1;
    * $A_1, \ldots, A_n$neslučitelné $\rightarrow$ pst, že nastane alespoň 1 z těchto jevů je rovna součtu pstí těchto jevů;
      · $\Pr(A_1 \cup \cdots \cup A_n) = \Pr(A_1) + \cdots + \Pr(A_n)$
      · Pr(liché číslo) = $\Pr(1 \cup 3 \cup 5) = 0.5$
      · Pr(padne 1) + Pr(padne 3) + Pr(padne 5) = $1/6 + 1/6 + 1/6 = 3/6 = 0.5$

* A_1 a A_2 neslučitelné $\rightarrow \Pr(A_1 \cup A_2) = \Pr(A_1) + \Pr(A_2)$
    * $\Pr(\text{padne } 1) = 1/6; \Pr(\text{padne } 2) = 1/6; \Pr(1 \cup 2) = 2/6$
    * Je-li jev A' opačný k jevu A, pak součet pstí těchto dvou jevů je 1
    * $\Pr(A) + \Pr(A') = 1$
        · Jev A ... padne sudé číslo $\Pr(A) = 0.5$
        · Jev A' ... padne liché číslo $\Pr(A') = 0.5$
        · Jev: buď padne sudé číslo nebo padne liché číslo: $\Pr(A \cup A') = 1$

Náhodné veličiny

- *náhodná veličina* X = pravidlo, které zobrazuje základní prostor možných výsledků do množiny reálních čísel

- i-tá realizace náh. veličiny se značí x_i

  (a) X ... počet pacientů v ordinaci za den $x_1 = 12$, $x_2 = 8$ ...
  (b) Y ... počet puntíků na vrchní straně kostky: $y_1 = 4$, $y_2 = 1$ ...
  (c) X... výška člověka v cm; $x_1 = 165\,cm$, $x_2 = 173\,cm$ ...
  (d) Y ... váha člověka v kg; $y_1 = 63\,kg$, $y_2 = 77\,kg$ ...

- Různá data $\rightarrow$ různé typy náhodných veličin

- 2 základní typy n.v.
  - diskrétní a), b)
  - spojité c), d)

3.2 Diskrétní náhodné veličiny

- nabývají konkrétních (převážně celých) hodnot
  - hod kostkou: padne 1,2,3,4,5,6. Nemůže padnout 3.5
  - $\Pr(X = 1) = 1/6$; $P(X = 1 \vee X = 2) = 1/3$

- pstní fce $p(x)$:
  - $p(x) = \Pr(X = x)$
  - $p(x)$ ... pst, že náh.veličina X se realizuje v hodnotě x
  - $p(x) \geq 0$, $\sum_{i=1}^{n} p(x_i) = 1$
  - pstní fce pro případ hod kostkou:

- distribuční funkce $F(x)$

- $F(x) = \Pr(X \leq x)$
- $F(x)$ ... pst že realizace náh.veličiny X nepřekročí hodnotu x
- příklad: distribuční fce hodu kostkou:

- $\Pr(\mathbf{X} > \mathbf{x}) = \mathbf{1} - \Pr(\mathbf{X} \leq \mathbf{x}) = \mathbf{1} - \mathbf{F}(\mathbf{x})$ ... DŮLEŽITÉÉÉÉ
- shrnutí:
  * $p(x) = P(X = x)$
  * $F(x) = \Pr(X \leq x) = \sum_{t=-\infty}^{x} p(t)$

- diskrétní náhodné veličiny se řídí nějakým diskrétním rozdělením:

1. Alternativní $A(\vartheta)$;
2. Binomické $Bi(n, \vartheta)$;
3. Geometrické $Geom(\vartheta)$;
4. Hypergeometrické $Hyper(N, M, k)$;
5. Poissonovo $Po(\lambda)$

Kombinační číslo

- $\binom{n}{k} = \dfrac{n!}{k!(n-k)!}$
- $\binom{5}{2} = \dfrac{5!}{2!(5-2)!} = \dfrac{5\,4\,3\,2\,1}{2\,1\,3\,2\,1} = \dfrac{20}{2} = 10$

```
choose(5,2)
## 10
```

Binomické rozdělení

- X popisuje počet úspěchů v posloupnosti n opakovaných nezávislých pokusů, přičemž pst úspěchu je v každém pokusu θ.
- pstní fce:

$$p(x) = \binom{n}{x} \theta^x (1-\theta)^{n-x} \text{ pro } x = 0, \dots, n$$

- $\theta \in \langle 0\,;\, 1 \rangle$, $E(X) = n\theta$; rozptyl: $D(X) = n\theta(1-\theta)$
- $p(x)$ ... dbinom(x, N, θ)
- $F(x)$ ... pbinom(x, N, θ)

Příklad 3.1. Pojišťovna zjistila, že 12 % pojistných událostí je způsobeno vloupáním. Jaká je pravděpodobnost, že mezi 30 náhodně vybranými pojistnými událostmi bude způsobeno vloupáním

a) nejvýše 6;

b) aspoň 6;

c) právě 6;

d) od dvou do pěti?

ad a)
```
sum(dbinom(0:6, size=30, prob=0.12))
pbinom(6, size=30, prob=0.12)
## [1] 0.9393926
```

ad b)
```
1-sum(dbinom(0:5, size=30, prob=0.12))
1-pbinom(5, size=30, prob=0.12)
## [1] 0.1430769
```

ad c)
```
dbinom(6, size=30, prob=0.12)
## [1] 0.08246953
```

ad d)
```
pbinom(5, size=30, prob=0.12) - pbinom(1, 30, 0.12)
sum(dbinom(2:5, size=30, prob=0.12))
## [1] 0.7469528
```

Geometrické rozdělení

- X ... popisuje počet neúspěchů před prvním úspěchem v posloupnosti nezávislých pokusů.
- pstní fce:
$$p(x) = (1-\theta)^x\theta \tag{1}$$
- $\theta \in \langle 0\,;\,1\rangle$, $E(X) = \dfrac{1-\theta}{\theta}$, $D(X) = \dfrac{1-\theta}{\theta^2}$
- $p(x)$ ... dgeom(x, θ)
- $F(x)$ ... pgeom(x, θ)

Příklad 3.2. Jaká je pravděpodobnost, že při hře Člověče, nezlob se! nasadíme figurku nejpozději při třetím hodu?

Počet neúspěchů: $x = 0, 1, 2$; pravděpodobnost úspěchu: $\theta = \frac{1}{6}$;

```
sum(dgeom(0:2, prob=1/6))
pgeom(2, prob=1/6)
## [1] 0.4212963
```

Hypergeometrické rozdělení

- Máme N objektů, mezi nimi je M objektů majících sledovanou vlastnost, $0 \leq M \leq N$. Náhodně bez vracení vybereme k objektů ($0 \leq k \leq N$).
- X popisuje, kolik z k vybraných objektů má sledovanou vlastnost.
- pstní fce
$$P_{N,M,k}(x) = \frac{\binom{M}{x}\binom{N-M}{k-x}}{\binom{N}{k}}. \tag{2}$$
- $p(x)$ ... dhyper(x, M, N-M, k)
- $F(x)$ ... phyper(x, M, N-M, k)

Příklad 3.3. Koupili jsme 10 cibulek červených tulipánů a 5 cibulek žlutých tulipánů. Zasadili jsme 8 náhodně vybraných cibulek.

a) Jaká je pravděpodobnost, že žádná nebude cibulka žlutých tulipánů?

b) Jaká je pravděpodobnost, že jsme zasadili všech 5 cibulek žlutých tulipánů?

c) Jaká je pravděpodobnost, že aspoň dvě budou cibulky žlutých tulipánů?

Počet objektů: $N = 15$, počet označených objektů: $M = 5$, počet vybraných objektů: $n = 8$

ad a)
```
dhyper(0, m=5, n=10, k=8)
## [1] 0.006993007
```

ad b)
```
dhyper(5, m=5, n=10, k=8)
## [1] 0.01864802
```

ad c)
```
sum(dhyper(2:5, m=5, n=10, k=8))
phyper(5, m=5, n=10, k=8)-phyper(1, m=5, n=10, k=8)
## [1] 0.8997669
```