4 Náhodné veličiny

- Víc než výsledek nás často zajímají jeho číselné interpretace
- *náhodná veličina* X je pravidlo, které zobrazuje základní prostor možných výsledků do množiny reálních čísel
- i-tá realizace náh. veličiny se značí x_i
  (a) $X \ldots$ výška člověka v cm; $x_1 = 165\, cm$, $x_2 = 173\, cm \ldots$
  (b) $Y \ldots$ váha člověka v kg; $y_1 = 63\, kg$, $y_2 = 77\, kg \ldots$
  (c) $Y \ldots$ počet pacientů v ordinaci za den $y_1 = 12$, $y_2 = 8 \ldots$
  (d) $X \ldots$ počet puntíků na vrchní straně kostky: $x_1 = 4$, $x_2 = 1 \ldots$
- s jakou pstí nabývá náh. veličina X určité hodnoty, nebo je obsažena v určitém intervalu hodnot: píšeme: $\Pr(X = x)$
- náhodné veličiny máme:
  - diskrétní - nabývají konkrétních (převážně celých) hodnot c), d)
    - hod kostkou: padne 1,2,3,4,5,6.
    - nemůže padnout 3.5
    - $\Pr(X = 4) = 1/6$
  - spojité - nabývají lib. hodnoty z daného intervalu a), b)
    - změříme výšku člověka:
      - základní prostor rozdělíme na intervaly: I1:0-100; I2:100-125; I3:125-150; I4:150-175; I5:175-200; I6:200-225;
      - pst, že naměřená hodnota bude náležet do intrvalu I4: 150-175cm
      - $\Pr(X \in I4) = \ldots$
- pstní fce $p(x)$ (X je diskrétní):
  - $p(x) = \Pr(X = x)$
  - pstní fce v bodě x je rovna psti, že náh.veličina X se realizuje v hodnotě x
  - pstní fce je nezáporná $\Pr(x) \geq 0$ a normovaná $\sum_{i=1}^{\infty} \Pr(X = x_i) = 1$
  - pstní fce pro případ hod kostkou:

- hustota $f(x)$ (X je spojitá)

- pst realizace X v libovolném intervalu I se dá vyjádřit jako plocha pod křivkou pomocí integrálního tvaru:
$$\Pr(X \in I) = \int_{x \in I} f(x)dx, \tag{1}$$
kde $f(x)$ je *hustota* pravděpodobnosti spojité náhodné veličiny
- hustota je nezáporná a normovaná (plocha pod křivkou hustoty = 1)
- příklad hustoty: Gaussova křivka

• distribuční funkce $F(x)$ (X je diskrétní nebo spojitá)
  - distr.fce v bodě x je rovna psti, že realizace náh.veličiny X nepřekročí hodnotu x
  - $F(x) = \Pr(X \leq x)$
  - příklad: distribuční fce hodu kostkou:

  - vlastnosti: neklesající, zprava spojitá, normovaná,
  - platí: $\Pr(X > x) = 1 - \Pr(X \leq x) = 1 - F(x)$ KOMPLEMENTARITA
  - platí v diskrétním případě:
    * $F(x) = \Pr(X \leq x) = \sum_{t=-\infty}^{x} p(t)$
    * $p(x) = P(X = x)$
  - platí ve spojitém případě:
    * $F(x) = \Pr(X \leq x) = \int_{-\infty}^{x} (t)dt$
    * $P(X = x) = 0$

Typy rozdělení

- máme náhodný výběr a rádi bychom věděli z jakého pochází rozložení.
- každé rozložení má své specifické parametry
  - diskrétní
    - Alternativní $A(\theta)$;
    - Binomické $\mathrm{Bin}(n, \theta)$;
    - Poissonovo $\mathrm{Po}(\lambda)$
    - Geometrické
    - Hypergeometrické
  - spojité
    - Rovnoměrné spojité
    - Exponenciální $Ex(\lambda)$
    - Normální $N(\mu, \sigma^2)$
    - Standardizované normální $N(0, 1)$
      - Pearsonovo $\chi^2(n)$
      - Studentovo $t(n)$
      - Fisherovo-Snedecorovo $F(n_1, n_2)$
    - dvourozměrné normální $N_2(\boldsymbol{\mu}, \boldsymbol{\Sigma})$
- nejdříve musíme odhadnout typ rozdělení, potom parametry rozdělení

Kombinační číslo

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

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

5 Binomické rozdělení $\mathrm{Bin}(n, \theta)$

- $X \sim \mathrm{Bin}(n, \theta)$
- $X \ldots$ počet úspěchů v posloupnosti n nezávislých opakovaných pokusů, přičemž pravděpodobnost úspěchu v každém pokusu je vyjádřena parametrem θ.
- pravděpodobnostní funcke:

$$p(x) = \begin{cases} \binom{n}{x}\theta^x(1-\theta)^{n-x} & \text{pro } x = 0, \ldots, n; \\ 0 & \text{jinak.} \end{cases}$$

- distribuční funkce:

$$F(x) = \sum_{t=0}^{x} \binom{n}{t} \theta^t (1-\theta)^{n-t}.$$

- vlastnosti: $E(X) = n\theta$; rozptyl: $D(X) = n\theta(1-\theta)$
- dbinom(x, n, θ), pbinom(x, n, θ)

Příklad 5.1. Předpokládejme, že pravděpodobnost výskytu dermatoglifického vzoru *vír* na palci pravé ruky mužů české populace, Pr(vír) = 0.533. Pravděpodobnost výskytu ostatních vzorů na palci pravé ruky u mužů potom bude Pr(ostatní) =......................... . Vypočítejte, jaká je pravděpodobnost, že mezi 10 muži bude výskyt vzoru *vír*

a. právě u šesti mužů;
b. nejvýše u šesti mužů;
c. alespoň u šesti mužů;
d. u dvou, tří, čtyř nebo pěti mužů.

X ....
Počet pokusů: n = ..., pravděpodobnost úspěchu: θ = ...

ad a. ....

....

```
dbinom(6, 10, 0.533)
```

S pravděpodobností .... % bude výskyt vzoru *vír* právě u šesti mužů z deseti.

ad b. ....

....

```
sum(dbinom(0:6, size=10, prob=0.533))
pbinom(6, size=10, prob=0.533)
```

S pravděpodobností .... % bude výskyt vzoru *vír* nejvýše u šesti mužů z deseti.

ad c. ....

....

```
1-sum(dbinom(0:5, size=10, prob=0.533))
1-pbinom(5, size=10, prob=0.533)
```

S pravděpodobností .... % bude výskyt vzoru *vír* alespoň u šesti mužů z deseti.

ad d. ....

....

```
pbinom(5, size=10, prob=0.533) - pbinom(1, 10, 0.533)
sum(dbinom(2:5, size=10, prob=0.533))
```

S pravděpodobností .... % bude výskyt vzoru *vír* u dvou, tří, čtyř nebo pěti mužů z deseti.

5.1 Poissonovo rozdělení Po(λ)

- $X \sim Po(\lambda)$
- Náhodná veličina X udává počet událostí, které nastanou v jednotkovém časovém intervalu, přičemž k událostem dochází náhodně, jednotlivě a vzájemně nezávisle. Parametr $\lambda > 0$ je střední počet těchto událostí. Píšeme $X \sim \mathrm{Po}(\lambda)$.

- pravděpodobnostní funkce:

$$p(x) = \begin{cases} \frac{\lambda^x}{x!}e^{-\lambda} & \text{pro x=0,1,...;} \\ 0 & \text{jinak.} \end{cases} \tag{2}$$

- distribuční funkce:

$$F(x) = \sum_{t=0}^{x} \frac{\lambda^t}{t!}e^{-\lambda}. \tag{3}$$

- vlastnosti: $E(X) = \lambda$; rozptyl: $D(X) = \lambda$
- dpois(x, lambda), ppois(x, lambda)

Příklad 5.2. Při provozu balicího automatu vznikají během směny náhodné poruchy, které se řídí rozdělením Po(2). Jaká je pravděpodobnost, že během směny dojde k alespoň jedné poruše?

X ... počet poruch během směny; $X \sim \mathrm{Po}(\lambda = \text{............})$;

$\Pr(X \geq 1) = 1 - \Pr(X < 1) = 1 - \Pr(X \leq 0) = 1 - \Pr(X = 0) = \text{..................................} = \text{.............}$

```
1-dpois(0, lambda=2)
1-ppois(0, lambda=2)
```

Pravděpodobnost, že během směny dojde k alespoň jedné poruše je .... %.

Příklad 5.3. Telefonní ústředna zapojí během hodiny průměrně 15 hovorů. Jaká je pravděpodobnost, že během 4 minut ústředna zapojí

a. právě jeden hovor?

```
#1h=60min .. 15 hovoru 4 min ... 1 hovor
dpois(1, 1)
```

b. alespoň dva hovory?

```
#1h=60min .. 15 hovoru 4 min ... 1 hovor
1-ppois(1,1)
```

c. nejméně tři a nejvýše čtyři hovory?

```
#1h=60min .. 15 hovoru 4 min ... 1 hovor
sum(dpois(3:4,1))
ppois(4,1)-ppois(2,1)
```

d. nejvýše pět hovorů?

```
#1h=60min .. 15 hovoru 4 min ... 1 hovor
ppois(5,1)
```

Pravděpodobnost, že během 4 minut zapojí ústředna právě jeden hovor je .... %. Pravděpodobnost, že během 4 minut zapojí ústředna právě alespoň dva hovory je .... %. Pravděpodobnost, že během 4 minut zapojí ústředna nejméně tři a nejvýše čtyři hovory je .... %. Pravděpodobnost, že během 4 minut zapojí ústředna právě nejvýše pět hovorů je .... %.