

# Domácí úkol na BI-PST 2018

Pavel Jahoda a Jan Lidák

December 10, 2018

## 1 Úkol 1

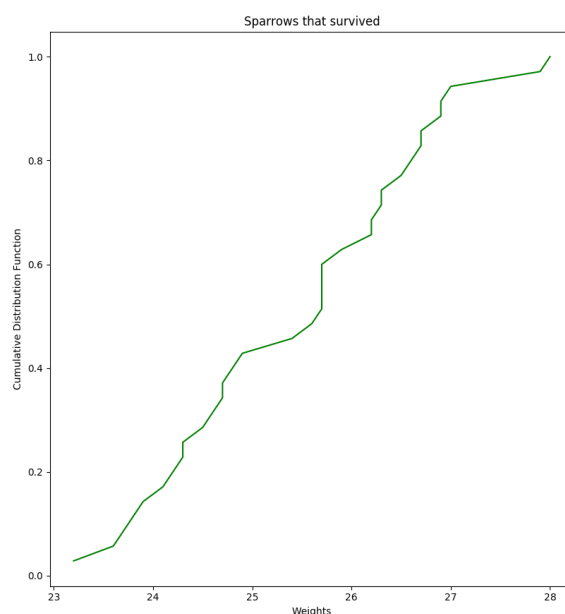
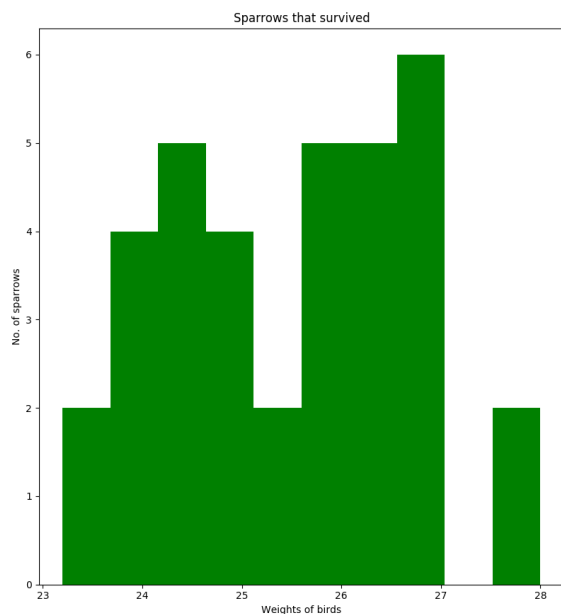
Data jsou z pozorování 59 vrabců během zimy. První veličina  $\mathbf{X}$  reprezentuje váhy vrabců v gramech. Druhá veličina  $\mathbf{Y}$  nabývá dvou hodnot 'survived', pokud vrabec přežil a 'perished' pokud nepřežil. Sledovanou proměnnou  $X$  jsme rozdělili na dvě pozorované skupiny  $\mathbf{X1}$  (vrabci co přežili) a  $\mathbf{X2}$  (vrabci co nepřežili).

$EX1=25.463$ ,  $\text{var}(X1)=1.539$  a medián je 25.7.

$EX2=26.275$ ,  $\text{var}(X2)=2.078$  a medián je 26.

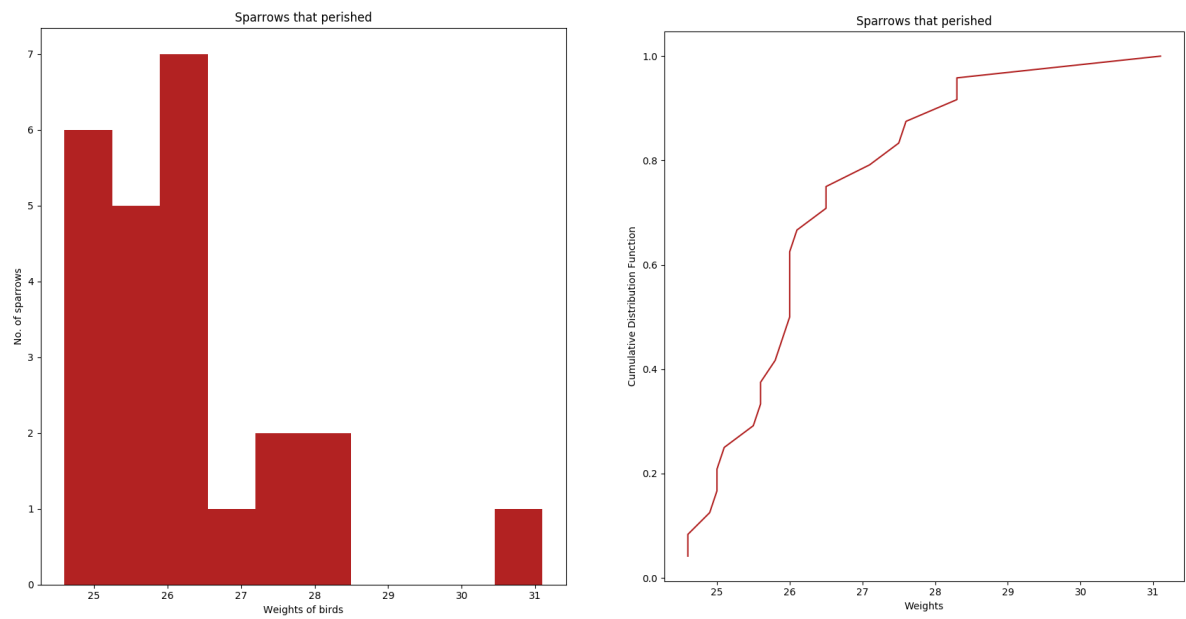
## 2 Úkol 2

Nejprve vykreslíme histogram a graf empirické distribuční funkce pro vrabce kteří přežili.



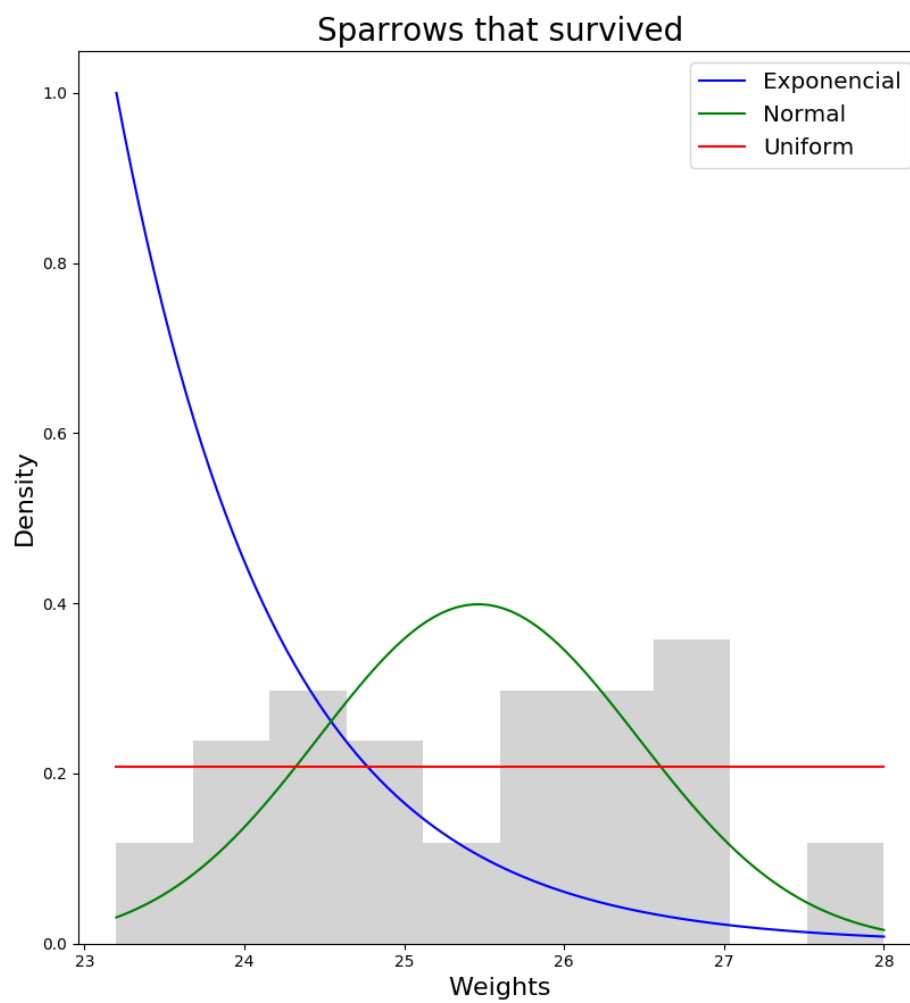
TODO.

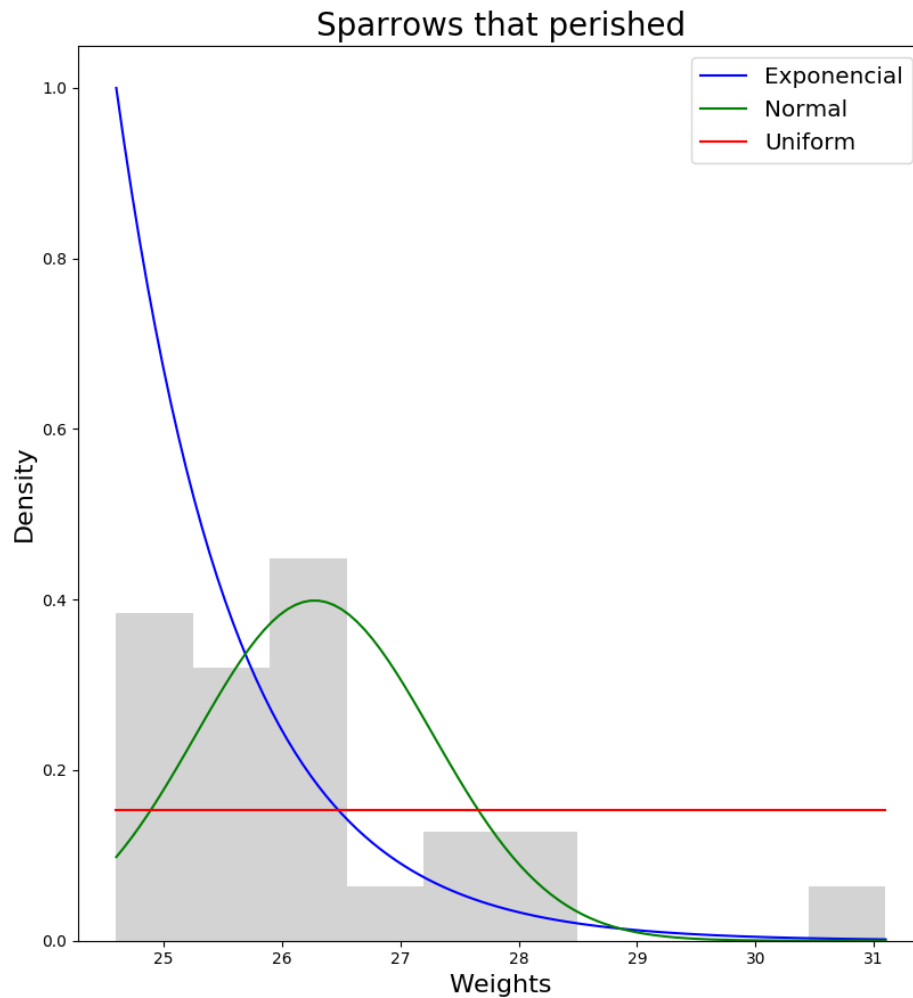
Poté vykreslíme histogram a graf empirické distribuční funkce pro vrabce kteří nepřežili.



Graf empirické distribuční funkce se podobá grafu exponenciálního rozdělení s parametrem  $\lambda = 1$ .

### 3 Úkol 3

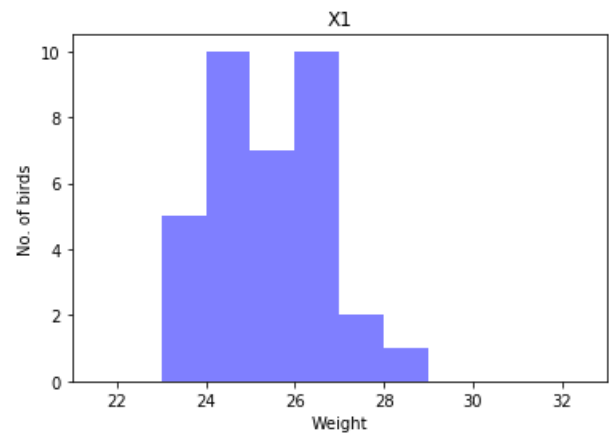
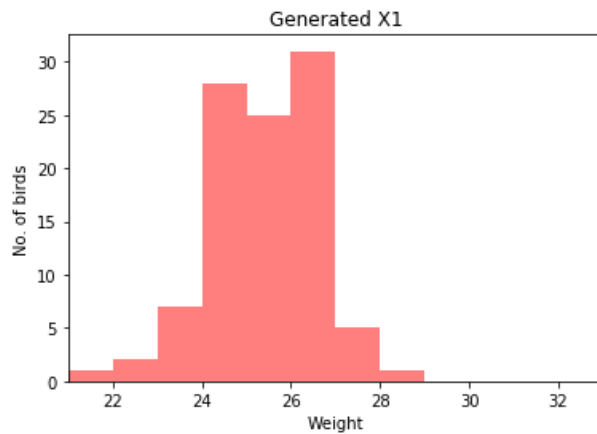




## 4 Úkol 4

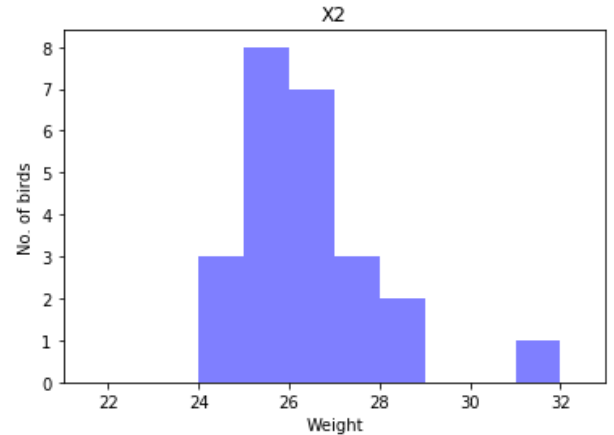
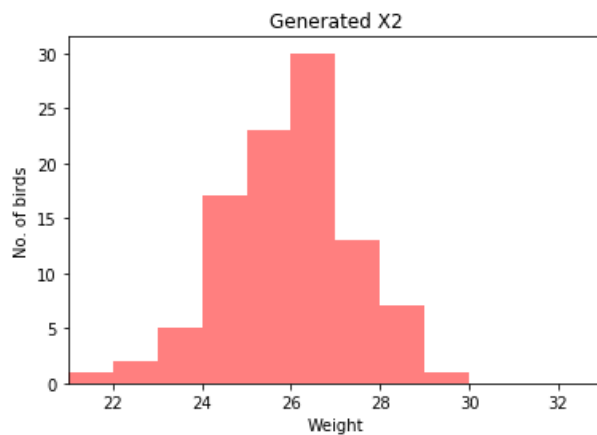
Na následujících grafech je zobrazeno 100 vygenerovaných hodnot spolu s načtenými daty z datasetu. Histogram přeživších opeřenců vypadá jako normálního rozdělení s parametry  $\lambda = EX = 25.793$  a  $\sigma^2 = \text{var}X = 1.918$ , hodnoty byli tedy generovány s těmito parametry.

První graf vygenerovaných dat poměrně silně připomíná získaná data, věříme tedy že toto rozdělení bylo zvoleno správně.



Data vygenerovaná pro mrtvé opěření už na tom nejsou tak dobře, část dat "utíká" do strany, zde by se hodilo mít větší dataset aby se dala lépe odhadnout distribuční funkce, popřípadě lépe vystihnout parametry rozdělení.

Pro opěření jež nepřežili byla data generována s normální distribuční funkcí s parametry  $\lambda = EX = 26.275$  a  $\sigma^2 = \text{var}X = 2.078$ .



## 5 Úkol 5

$\bar{X}_n$  je výběrový průměr veličiny,  $\sigma$  je směrodatná odchylka,  $n$  je počet prvků v rozdělení. Spolehlivost má být 95%, tedy

$$\alpha = 0.05 \Rightarrow \alpha/2 = 0.025$$

$$z_{\alpha/2} = z_{0.025} = \Phi^{-1}(0.975) \doteq 1.96$$

Oboustranný 95% konfidenční interval spočteme jako:

$$(S_{X_1}, U_{X_1}) = (\bar{X}_1 - z_{\alpha/2} \cdot \frac{\sigma_1}{\sqrt{n_1}}, \bar{X}_1 + z_{\alpha/2} \cdot \frac{\sigma_1}{\sqrt{n_1}})$$

$$= (25,463 - 1,96 \cdot \frac{1,24}{5,9}, 25,463 + 1,96 \cdot \frac{1,24}{5,9}) = (25.052, 25.874) \quad (1)$$

$$\begin{aligned}
(S_{X_2}, U_{X_2}) &= (\overline{X}2_n - z_{a/2} \cdot \frac{\sigma_2}{\sqrt{n_2}}, \overline{X}2_n + z_{a/2} \cdot \frac{\sigma_2}{\sqrt{n_2}}) \\
&= (26.275 - 1, 96 \cdot \frac{1,44}{4,9}, 26.275 + 1, 96 \cdot \frac{1,44}{4,9}) = (25.698, 26.852)
\end{aligned} \tag{2}$$