

Odhad rozdělení latence odezvy neuronu

ZBYNĚK PAWLAS

Katedra pravděpodobnosti a matematické statistiky
Matematicko-fyzikální fakulta
Univerzita Karlova v Praze

17. letní škola JČMF ROBUST 2012
10. září 2012, Němčičky

Zadání úlohy

nezávislé náhodné výběry:

- X_1, \dots, X_n náhodný výběr z rozdělení s distrib. funkcí F
- Y_1, \dots, Y_n náhodný výběr z rozdělení s distrib. funkcí G
- $\tilde{Y}_1, \dots, \tilde{Y}_m$ náhodný výběr z rozdělení s distrib. funkcí G

Data:

$$Z_i = \min(X_i, Y_i), \quad i = 1, \dots, n$$

$$\tilde{Y}_j, \quad j = 1, \dots, m$$

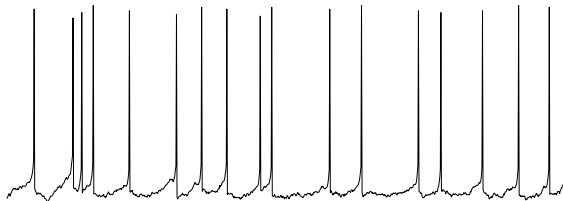
Cíl: odhad F

Motivace z neurofyzologie

informace v nervovém systému je přenášena posloupností akčních potenciálů (tzv. **spiků**) generovaných jednotlivými neurony

pro přenos signálu je daleko důležitější doba než tvar nebo amplituda spiku → **časové kódování**

dobu spiku lze považovat za jeden bod na časové ose



Motivace z neurofyzologie

za přítomnosti spontánní aktivity jsou generované *spontánní* spiky

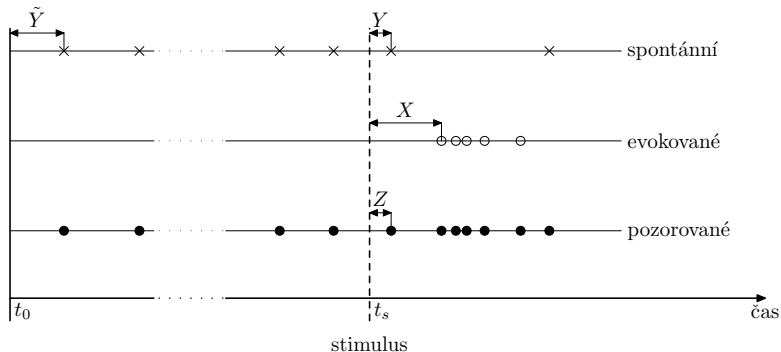
jejich časové uspořádání můžeme považovat za realizaci stacionárního bodového procesu

zajímá nás reakce neuronu na vnější stimulus (např. zvukový podnět), ta se projeví jedním nebo více *evokovanými* spiky

latence odezvy neuronu je doba od začátku stimulu po výskyt prvního evokovaného spiku

v praxi nejsme schopni rozlišit, jestli první pozorovaný spike po stimulu byl způsoben spontánní aktivitou nebo reakcí na podnět

Schématické znázornění



Značení a data

\check{Y} ... doba do výskytu prvního *spontánního* spiku za přítomnosti pouze spontánní aktivity

Y ... doba do výskytu prvního *spontánního* spiku po začátku stimulu

X ... doba do výskytu prvního *evokovaného* spiku po začátku stimulu

Z ... doba do výskytu prvního spiku po začátku stimulu

opakované stimulace za stejných podmínek

měření:

$$\check{Y}_1, \dots, \check{Y}_m, \quad Z_1, \dots, Z_n$$

Vztahy mezi distribučními funkcemi

nezáporné náhodné veličiny:

- X má distribuční funkci F
- Y i \tilde{Y} mají distribuční funkci G
- $Z = \min(X, Y)$ má distribuční funkci H

$$1 - H(t) = (1 - F(t))(1 - G(t))$$

$$F(t) = 1 - \frac{1 - H(t)}{1 - G(t)}, \quad t < \tau_G$$

$$\tau_G = \sup\{t \in \mathbb{R} : G(t) < 1\}, \quad \tau_H = \sup\{t \in \mathbb{R} : H(t) < 1\}$$

$$\tau_G \geq \tau_H$$

Neparametrický odhad

G i H můžeme odhadnout z dat příslušnými empirickými distribučními funkcemi:

$$G_m(t) = \frac{1}{m} \sum_{i=1}^m \mathbf{1}\{\tilde{Y}_i \leq t\},$$

$$H_n(t) = \frac{1}{n} \sum_{i=1}^n \mathbf{1}\{Z_i \leq t\},$$

pak přirozený neparametrický odhad F je

$$F_{n,m}(t) = 1 - \frac{1 - H_n(t)}{1 - G_m(t)}$$

Vlastnosti odhadu

vlastnosti funkce $F_{n,m}(t)$:

- je dobře definována pro $t < \max \tilde{Y}_i$, jinak položíme $F_{n,m}(t) = 1$,
- po částech spojitá funkce, nabývá skoky v bodech $\tilde{Y}_1, \dots, \tilde{Y}_m, Z_1, \dots, Z_n$,
- není nutně monotónní,
- může nabývat záporných hodnot

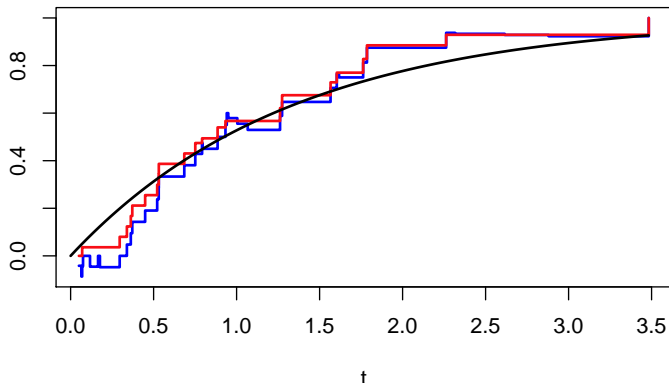
modifikace odhadu $F_{n,m}$:

$$\tilde{F}_{n,m}(t) = 1 - \frac{1 - \inf_{s \geq t} F_{n,m}(s)}{1 - \inf_{s \geq 0} F_{n,m}(s)}$$

Vykreslení odhadů pro simulovaná data

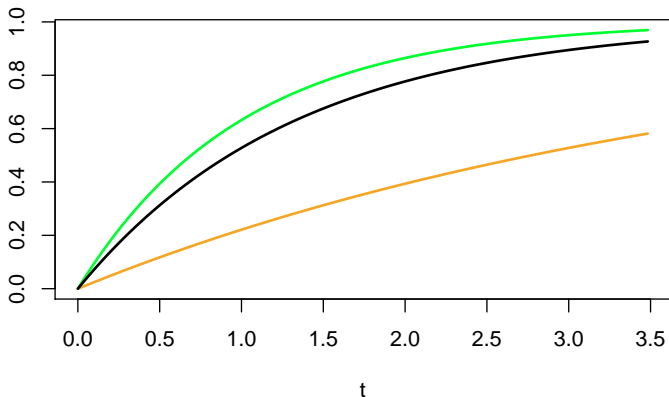
$Y \sim \text{Exp}(0,25)$, $Z \sim \text{Exp}(1)$, $n = m = 25$

$$F_{n,m}, \tilde{F}_{n,m}, F$$



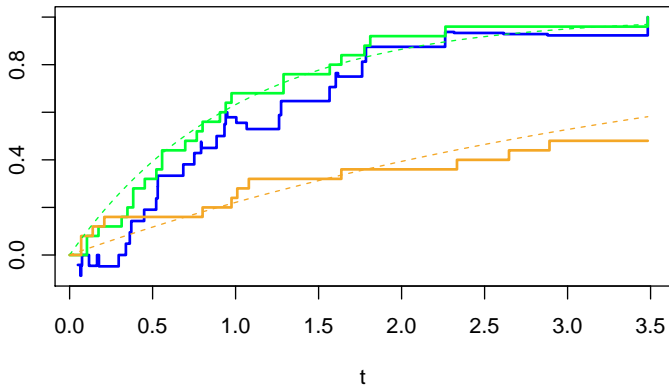
Teoretické distribuční funkce

H , G , F



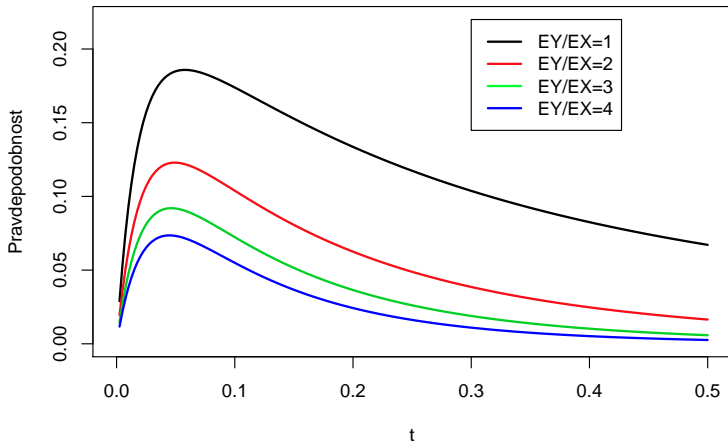
Empirické distribuční funkce

$$H_n, G_m, F$$



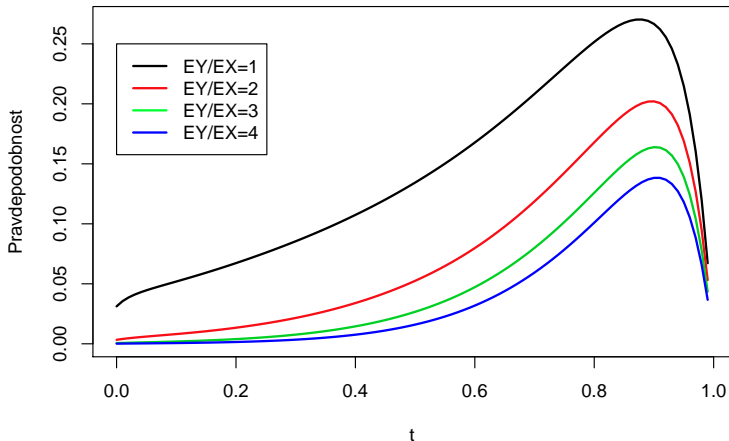
Graf $P(F_{n,m}(t) < 0)$

$Y \sim \text{Exp}, Z \sim \text{Exp}(1), n = m = 25$



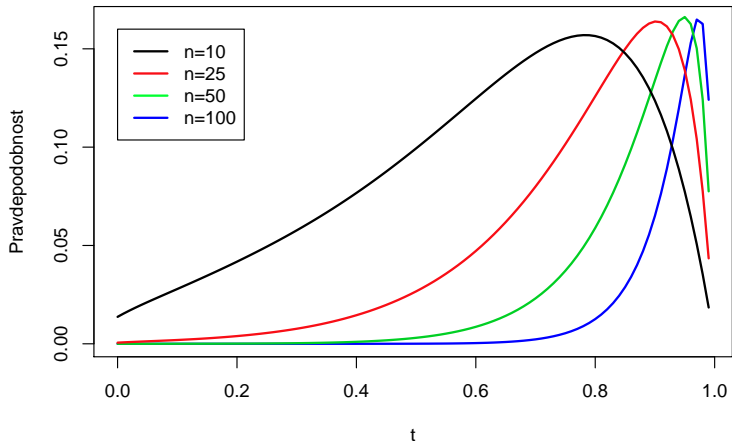
Graf $P(F_{n,m}(t) > F_{n,m}(1))$

$Y \sim \text{Exp}, Z \sim \text{Exp}(1), n = m = 25$



Graf $P(F_{n,m}(t) > F_{n,m}(1))$

$Y \sim \text{Exp}(0,25)$, $Z \sim \text{Exp}(1)$, $n = m$



Silná konzistence

pro každé $t < \tau_G$ plyne ze silného zákona velkých čísel:

$$F_{n,m}(t) \longrightarrow F(t) \quad \text{s.j. pro } n, m \rightarrow \infty$$

$$\tilde{F}_{n,m}(t) \longrightarrow F(t) \quad \text{s.j. pro } n, m \rightarrow \infty$$

Slabá konvergence empirického procesu

uvažujme $\delta < \tau_G$

$$Y_{n,m}(t) = \sqrt{n} (F_{n,m}(t) - F(t)), \quad t \in [0, \delta]$$

konverguje slabě pro $n, m \rightarrow \infty$ v $D[0, \delta]$ k centrovanému gaussovskému procesu

$$Y(t) = \frac{1}{1 - G(t)} W_H^0(t) - \frac{1 - F(t)}{1 - G(t)} W_G^0(t), \quad t \in [0, \delta]$$

kde $\{W_K^0(t) : t \in [0, \delta]\}$ je centrovaný gaussovský proces s kovarianční funkcí

$$\mathbb{E} W_K^0(s) W_K^0(t) = K(s \wedge t) - K(s)K(t)$$

Případ známého G

předpokládáme, že distribuční funkce G je známá (odpovídá $m = \infty$)

$$F_{n,\infty}(t) = 1 - \frac{1 - H_n(t)}{1 - G(t)}$$

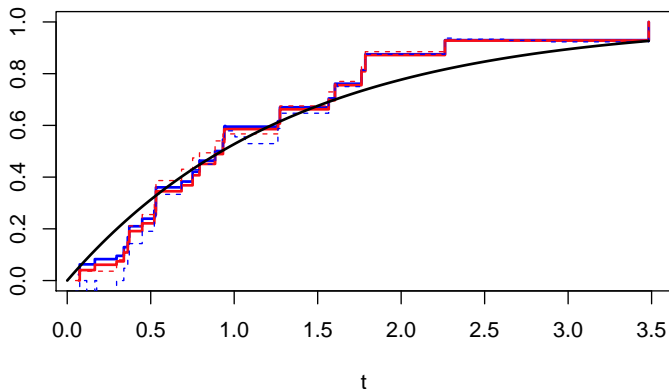
- není nutně monotónní
- může nabývat záporných hodnot

$$\begin{aligned} P(F_{n,\infty}(t) < 0) &= P(H_n(t) < G(t)) \\ &= \sum_{k < nG(t)} \binom{n}{k} H(t)^k (1 - H(t))^{n-k} \end{aligned}$$

Vykreslení odhadů pro simulovaná data

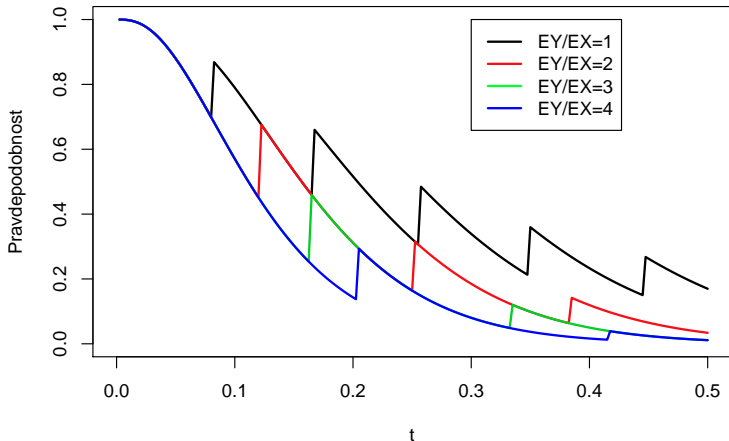
$Y \sim \text{Exp}(0,25)$, $Z \sim \text{Exp}(1)$, $n = 25$, $m = \infty$

$F_{n,\infty}$, $F_{n,n}$, $\tilde{F}_{n,\infty}$, $\tilde{F}_{n,n}$, F



Graf $P(F_{n,\infty}(t) < 0)$

$Y \sim \text{Exp}, Z \sim \text{Exp}(1), n = 25$



Exponenciální rozdělení

v motivačním příkladu: spontánní aktivita se řídí podle Poissonova procesu

$$G(t) = P(Y \leq t) = 1 - e^{-\lambda t}$$

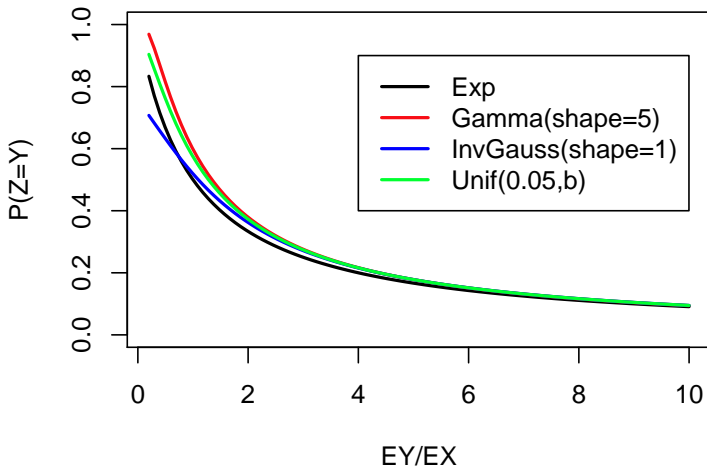
$$\hat{F}(t) = 1 - (1 - H_n(t)) \exp\{\hat{\lambda}t\}$$

pravděpodobnost, že první spike bude spontánní:

$$P(Y \leq X) = P(Z = Y) = \frac{EZ}{EY}$$

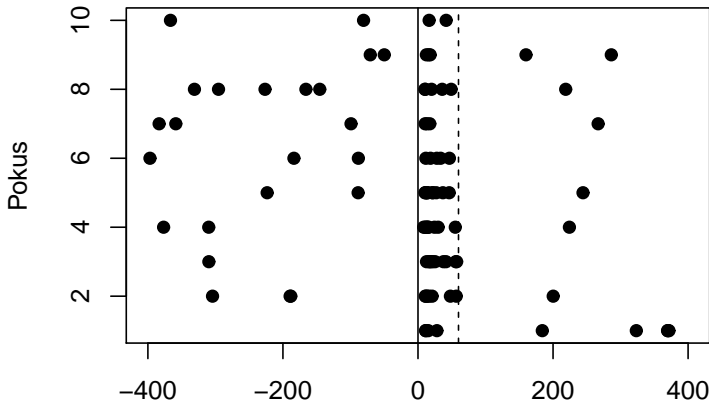
lze odhadnout z dat

Pravděpodobnost, že první spike je spontánní

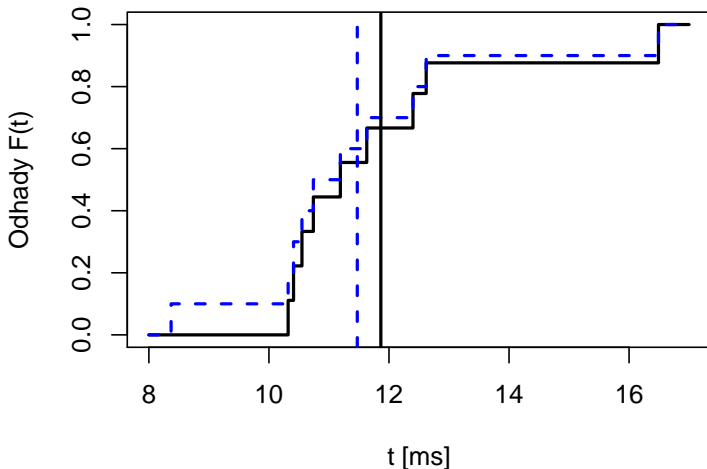


Experimentální data

nervová buňka sluchového systému
 zvukový stimulus o intenzitě 40 dB



Odhady distribuční funkce a střední hodnoty



Simultánní záznam několika neuronů

N neuronů zaznamenáváno simultánně

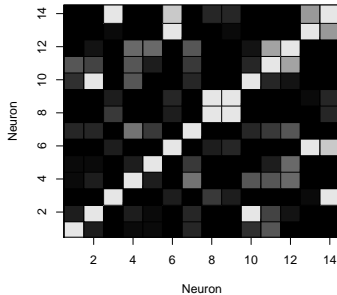
pro dvojici neuronů: $(X^{(1)}, X^{(2)}), (Y^{(1)}, Y^{(2)})$,
 $Z^{(i)} = \min(X^{(i)}, Y^{(i)})$

$$P(X^{(1)} > s, X^{(2)} > t) = \frac{P(Z^{(1)} > s, Z^{(2)} > t)}{P(Y^{(1)} > s, Y^{(2)} > t)}$$

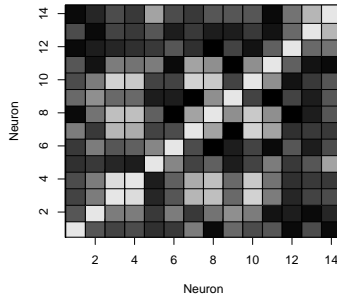
odhad nahrazením empirickými protějšky

$$\text{corr}(X^{(1)}, X^{(2)}) = \frac{\int_0^\infty \int_0^\infty P(X^{(1)} > s, X^{(2)} > t) ds dt - \mathbb{E}X^{(1)}\mathbb{E}X^{(2)}}{\sqrt{\text{var } X^{(1)}}\sqrt{\text{var } X^{(2)}}}$$

Odhad korelací latence odezvy v reálných datech



40 dB



80 dB

Zpoždění v odpovědi

evokovaný spike nemůže nastat před $t_s + \theta$
 $\theta \geq 0$ neznámý parametr

$$F(t) = 0 \quad \text{pro } t \in [0, \theta]$$

$$X = \theta + X^*$$

$$F(t) = F^*(t - \theta)$$

$$H(t) = G(t) + F^*(t - \theta)(1 - G(t))\mathbf{1}\{t > \theta\}$$

$$\theta = \inf\{t > 0 : H(t) \neq G(t)\}$$

Dekonvoluce

X, Y nezávislé náhodné veličiny
pozorujeme $Z = X + Y$
konvoluce:

$$H = F * G$$

zajímá nás F

Laplaceova transformace:

$$L_H(s) = L_F(s)L_G(s)$$

$$L_F(s) = \frac{L_H(s)}{L_G(s)}$$

Literatura



Z. PAWLAS, L. B. KLEBANOV, V. BENEŠ, M. PROKEŠOVÁ,
J. POPELÁŘ, P. LÁNSKÝ (2010):
First-spike latency in the presence of spontaneous activity.
Neural Computation **22**, 1675–1697.

Děkuji za pozornost