Machine Learning Worksheet 2

Thomas Blocher – MatrNr. 03624034
Raphael Dümig – MatrNr. 03623199

October 26, 2015

Problem 1

- $T \in \{0,1\}$: the scanned person is a civilian or terrorist respectively
- $S \in \{0,1\}$: scanner identifies the person as a civilian or terrorist respectively

$$\begin{aligned} P(S=1|T=1) &= 0.95 \\ P(S=0|T=0) &= 0.95 \\ P(T=1) &= 0.01 \end{aligned}$$

$$\begin{aligned} \Rightarrow P(T=0) &= 1 - P(T=1) = 0.99 \\ \Rightarrow P(S=1|T=0) &= 1 - P(S=0|T=0) = 0.05 \end{aligned}$$

$$\begin{aligned} P(T=1|S=1) &= \frac{P(S=1|T=1) \cdot P(T=1)}{P(S=1)} \\ &= \frac{P(S=1|T=1) \cdot P(T=1)}{\sum\limits_{t \in \{0,1\}} P(S=1|T=t) \cdot P(T=t)} \\ &= \frac{0.95 \cdot 0.01}{0.05 \cdot 0.99 + 0.95 \cdot 0.01} \\ &= \frac{1}{6} \end{aligned}$$

Problem 2

- $C = \{0, 1\}$: the tossed coin showes head or tail respectively

$$C \sim \text{Ber}\left(\frac{1}{2}\right)$$

- $T = \{0, 1, 2\}$: the number of two coin tosses that resulted in *tail*

$$\Rightarrow T \sim \text{Bin}\left(2, \frac{1}{2}\right)$$

- $R = \{0, 1, 2, 3\}$: number of red balls drawn from the box

$$R \sim \text{Bin}\left(3, \frac{t}{2}\right)$$

$$\begin{aligned}
P(T = t) &= \binom{2}{t} \cdot \left(\frac{1}{2}\right)^t \cdot \left(\frac{1}{2}\right)^{2-t} = \frac{1}{4} \cdot \binom{2}{t} \\
P(R = r|T = t) &= \binom{3}{r} \cdot \left(\frac{t}{2}\right)^r \cdot \left(1 - \frac{t}{2}\right)^{3-r} \\
P(T|R) &= \frac{P(R|T) \cdot P(T)}{P(R)}
\end{aligned}$$

$$\begin{aligned}
\Rightarrow P(T = 2|R = 3) &= \frac{P(R = 3|T = 2) \cdot P(T = 2)}{P(R = 3)} \\
&= P(R = 3|T = 2) \cdot \frac{P(T = 2)}{P(R = 3)} \\
&= P(R = 3|T = 2) \cdot \frac{P(T = 2)}{\sum_{t \in T} P(R = 3|T = t) \cdot P(T = t)} \\
&= 1 \cdot \frac{\frac{1}{4}}{0 \cdot \frac{1}{4} + \frac{1}{8} \cdot \frac{1}{2} + 1 \cdot \frac{1}{4}} \\
&= \frac{\frac{1}{4}}{\frac{5}{16}} = \frac{4}{5}
\end{aligned}$$

Problem 3

random variables:

- U: the selected urn

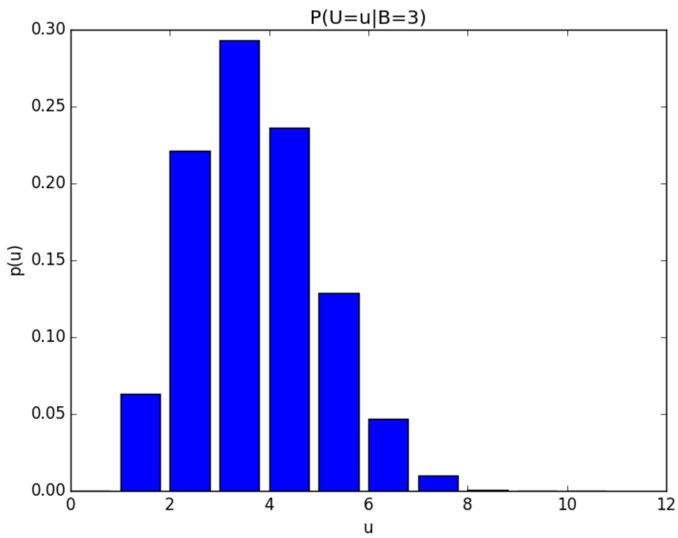
- B: the number of black balls drawn from an urn

$$\begin{aligned}
U &\sim \mathrm{U}(-1, 10) \\
B &\sim \mathrm{Bin}\left(10, 10 \cdot \frac{u}{10}\right) = \mathrm{Bin}(10, u) \\
\Rightarrow P(U) &= \frac{1}{11} \\
\Rightarrow P(B|U) &= \binom{10}{b} \cdot \left(\frac{u}{10}\right)^b \cdot \left(1 - \frac{u}{10}\right)^{10-b}
\end{aligned}$$

$$\begin{aligned}
P(U|B) &= \frac{P(B|U) \cdot P(U)}{P(B)} \\
P(B) &= \sum_{u' \in U} P(B|U = u') \cdot P(U = u') \\
&= \sum_{u' \in U} \binom{10}{b} \cdot \left(\frac{u'}{10}\right)^b \cdot \left(1 - \frac{u'}{10}\right)^{10-b} \cdot \frac{1}{11} \\
&= \frac{1}{11} \cdot \binom{10}{b} \cdot \sum_{u' \in U} \left(\frac{u'}{10}\right)^b \cdot \left(1 - \frac{u'}{10}\right)^{10-b} \\
&= \frac{1}{11} \cdot \binom{10}{b} \cdot \sum_{u'=0}^{10} \left(\frac{u'}{10}\right)^b \cdot \left(1 - \frac{u'}{10}\right)^{10-b}
\end{aligned}$$

$$\begin{aligned}
P(U = u|B = b) &= \frac{P(B = b|U = u) \cdot P(U = u)}{P(B = b)} \\
&= \frac{\binom{10}{b} \cdot \left(\frac{u}{10}\right)^b \cdot \left(1 - \frac{u}{10}\right)^{10-b} \cdot \frac{1}{11}}{\frac{1}{11} \cdot \binom{10}{b} \cdot \sum_{u'=0}^{10} \left(\frac{u'}{10}\right)^b \cdot \left(1 - \frac{u'}{10}\right)^{10-b}} \\
&= \frac{\left(\frac{u}{10}\right)^b \cdot \left(1 - \frac{u}{10}\right)^{10-b}}{\sum_{u'=0}^{10} \left(\frac{u'}{10}\right)^b \cdot \left(1 - \frac{u'}{10}\right)^{10-b}} \\
P(U = u|B = 3) &= \frac{\left(\frac{u}{10}\right)^3 \cdot \left(1 - \frac{u}{10}\right)^7}{\sum_{u'=0}^{10} \left(\frac{u'}{10}\right)^3 \cdot \left(1 - \frac{u'}{10}\right)^7}
\end{aligned}$$

$$\Rightarrow P(U=0|B=3) = 0$$
$$\Rightarrow P(U=1|B=3) = \frac{1594323}{25277575}$$
$$\Rightarrow P(U=2|B=3) = \frac{16777216}{75832725}$$
$$\Rightarrow P(U=3|B=3) = \frac{7411887}{25277575}$$
$$\Rightarrow P(U=4|B=3) = \frac{5971968}{25277575}$$
$$\Rightarrow P(U=5|B=3) = \frac{390625}{3033309}$$
$$\Rightarrow P(U=6|B=3) = \frac{1179648}{25277575}$$
$$\Rightarrow P(U=7|B=3) = \frac{250047}{25277575}$$
$$\Rightarrow P(U=8|B=3) = \frac{65536}{75832725}$$
$$\Rightarrow P(U=9|B=3) = \frac{243}{25277575}$$
$$\Rightarrow P(U=10|B=3) = 0$$



(result calculated using *ipython* interactive console – log attached to mail)

Problem 4

- $X = 1$: next ball drawn from one of the urns is black

- $X = 0$: next ball drawn from one of the urns is white

$$X \sim \mathrm{Ber}\left(\frac{u}{10}\right)$$

$$\begin{aligned}
P(X = 1) &= \sum_{u=0}^{10} P(X = 1 | U = u) \cdot P(U = u) \\
&= \frac{673471}{2022206} \\
&\approx 0.33304
\end{aligned}$$

(result calculated using *ipython* interactive console – log attached to mail)

Problem 5

$$p(x) = \begin{cases} 1, & x \in [0, 1] \\ 0, & x \notin [0, 1] \end{cases}$$

$$\begin{aligned}
\mathrm{E}[X] &= \int\limits_{-\infty}^{\infty} p(x) \cdot x \, \mathrm{d}x \\
&= \int\limits_{-\infty}^{0} p(x) \cdot x \, \mathrm{d}x + \int\limits_{0}^{1} p(x) \cdot x \, \mathrm{d}x + \int\limits_{1}^{\infty} p(x) \cdot x \, \mathrm{d}x \\
&= \int\limits_{-\infty}^{0} 0 \cdot x \, \mathrm{d}x + \int\limits_{0}^{1} 1 \cdot x \, \mathrm{d}x + \int\limits_{1}^{\infty} 0 \cdot x \, \mathrm{d}x \\
&= \int\limits_{-\infty}^{0} 0 \, \mathrm{d}x + \int\limits_{0}^{1} x \, \mathrm{d}x + \int\limits_{1}^{\infty} 0 \, \mathrm{d}x \\
&= 0 + \left[\frac{1}{2} \cdot x^2\right]_{x=0}^{1} + 0 \\
&= \frac{1}{2}
\end{aligned}$$

$$
\begin{aligned}
\mathrm{E}[X^2] &= \int_{-\infty}^{\infty} p(x) \cdot x^2 \, \mathrm{d}x = \int_0^1 x^2 \, \mathrm{d}x \\
&= \left[\frac{1}{3} \cdot x^3\right]_{x=0}^{1} \\
&= \frac{1}{3}
\end{aligned}
$$

$$
\begin{aligned}
\mathrm{Var}[X] &= \mathrm{E}[X^2] - \mathrm{E}[X]^2 \\
&= \frac{1}{3} - \left(\frac{1}{2}\right)^2 = \frac{1}{3} - \frac{1}{4} \\
&= \frac{1}{12}
\end{aligned}
$$

Problem 6

- $\mathrm{E}[X] = \mathrm{E}_Y[\mathrm{E}_{X|Y}[X]]$:

$$
\begin{aligned}
\mathrm{E}_Y[\mathrm{E}_{X|Y}[X]] &= \mathrm{E}_Y\left[\int \frac{p(x,y)}{p(y)} \cdot x \, \mathrm{d}x\right] \\
&= \int p(y) \cdot \int \frac{p(x,y)}{p(y)} \cdot x \, \mathrm{d}x \, \mathrm{d}y \\
&= \int p(y) \cdot \frac{1}{p(y)} \int p(x,y) \cdot x \, \mathrm{d}x \, \mathrm{d}y \\
&= \int \int p(x,y) \cdot x \, \mathrm{d}x \, \mathrm{d}y \\
&= \mathrm{E}[X]
\end{aligned}
$$

- $\mathrm{Var}[x] = \mathrm{E}_Y[\mathrm{Var}_{X|Y}[X]] - \mathrm{Var}_Y[\mathrm{E}_{X|Y}[X]]$:

$$
\begin{aligned}
\mathrm{E}_Y[\mathrm{Var}_{X|Y}[X]] - \mathrm{Var}_Y[\mathrm{E}_{X|Y}[X]] &= \mathrm{E}_Y[\mathrm{E}_{X|Y}[X^2] - \mathrm{E}_{X|Y}[X]^2] - \mathrm{Var}_Y[\mathrm{E}_{X|Y}[X]] \\
&= \mathrm{E}_Y[\mathrm{E}_{X|Y}[X^2] - \mathrm{E}_{X|Y}[X]^2] - \mathrm{E}_Y[\mathrm{E}_{X|Y}[X]^2] - \mathrm{E}_Y[\mathrm{E}_{X|Y}[X]]^2 \\
&= \mathrm{E}_Y[\mathrm{E}_{X|Y}[X^2]] - \mathrm{E}_Y[\mathrm{E}_{X|Y}[X]]^2 \\
&= \mathrm{E}[X^2] - \mathrm{E}[X]^2 \\
&= \mathrm{Var}[X]
\end{aligned}
$$

Problem 7

- $P[X > c] \leq \frac{\mathrm{E}[X]}{c}$:

$$\begin{aligned}
\frac{\mathrm{E}[X]}{c} &= \frac{\sum(p(x)\cdot x)}{c} \\
&= \frac{\sum_{x=-\inf}^{c-1}(p(x)\cdot x) + \sum_{x=c}^{\inf}(p(x)\cdot x)}{c} \\
&\geq \frac{\sum_{x=c}^{\inf}(p(x)\cdot x)}{c} \\
&\geq \frac{\sum_{x=c}^{\inf}(p(x)\cdot c)}{c} \\
&= \sum_{x=c}^{\inf}(p(x)) = P(X \geq c) \\
&\geq P(X > c)
\end{aligned}$$

- $P[X > \frac{3}{4}]$:

$$P[X > \frac{3}{4}] \leq \frac{\frac{n}{2}}{n * \frac{3}{4}}$$

Problem 8

- $P[|X - \mathrm{E}[x] > \alpha] \leq \frac{\mathrm{Var}[X]}{\alpha^2}$:

$$\begin{aligned}
\frac{\mathrm{Var}[X]}{\alpha} &= \frac{\mathrm{E}[(X-\mathrm{E}[x])^2]}{\alpha^2} \\
&\geq P((X-\mathrm{E}[x])^2 > \alpha^2) \\
&= P(|X-\mathrm{E}[x]| > \alpha)
\end{aligned}$$

- $P[|X| > \frac{3}{4}]$:

$$P[|X| > \frac{3}{4}] \leq \frac{n * \frac{1}{4}}{n * n * \frac{9}{12}}$$

$$P[X > \frac{3}{4}] \leq \frac{\frac{n*\frac{1}{4}}{n*n*\frac{9}{12}}}{2}$$

Problem 9

Basics

$\lambda Trick:$
$\sum_{i=0}^{n+1}\lambda_i = 1 \wedge \lambda'_i = \frac{\lambda_i}{(1-\lambda_{n+1})} \rightarrow \sum_{i=0}^{n}\lambda'_i = 1$

Definition Convex:
$\forall x_1, x_2 \forall t \in [0,1]\ f(tx_1 + (1-t)x_2) \leq t * f(x_1) + (1-t) * f(x_2)$

Induction

Hypothese

$f(\sum_{i=0}^{n} \lambda_i x_i) \leq \sum_{i=0}^{n} \lambda_i f(x_i) | \sum_{i=0}^{n} \lambda_i = 1$

Induction Start

$n = 1 \wedge \lambda_1 = 1$
$f(\lambda_1 x_1) \leq \lambda_1 * f(x_1)$
$f(1 * x_1) \leq 1 * f(x_1)$
$f(x_1) \leq f(x_1)$

Induction Step

$f(\sum_{i=0}^{n+1} \lambda_i x_i) \leq \sum_{i=0}^{n+1} \lambda_i f(x_i)$
Steps:
$f(\sum_{i=0}^{n+1} \lambda_i x_i) = f(\sum_{i=0}^{n} \lambda_i x_i + \lambda_{n+1} x_{n+1})$
$f(\sum_{i=0}^{n} \lambda_i x_i + \lambda_{n+1} x_{n+1}) = f((1-\lambda_{n+1}) \sum_{i=0}^{n} \lambda_i' x_i + \lambda_{n+1} x_{n+1})$
$f((1-\lambda_{n+1}) \sum_{i=0}^{n} \lambda_i' x_i + \lambda_{n+1} x_{n+1}) \leq (1-\lambda_{n+1}) f(\sum_{i=0}^{n} \lambda_i' x_i) + \lambda_{n+1} f(x_{n+1})$
$(1-\lambda_{n+1}) f(\sum_{i=0}^{n} \lambda_i' x_i) + \lambda_{n+1} f(x_{n+1}) \leq (1-\lambda_{n+1}) \sum_{i=0}^{n} \lambda_i' f(x_i) + \lambda_{n+1} f(x_{n+1})$
$(1-\lambda_{n+1}) \sum_{i=0}^{n} \lambda_i' f(x_i) + \lambda_{n+1} f(x_{n+1}) = \sum_{i=0}^{n} \lambda_i f(x_i) + \lambda_{n+1} f(x_{n+1})$
$\sum_{i=0}^{n} \lambda_i f(x_i) + \lambda_{n+1} f(x_{n+1}) = \sum_{i=0}^{n+1} \lambda_i f(x_i)$