Machine Learning Worksheet 4

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

November 8, 2015

Problem 1

Since we throw the coin until one Head occurce E(H) =1

The number of Tails: expected:
$E(H) = \sum_{n=0}^{\text{inf}} (\frac{1}{2}^n * n * \frac{1}{2})$
$= \frac{1}{2} * \sum_{n=0}^{\text{inf}} (\frac{1}{2}^n * n) = \frac{1}{2} * 2 = 1$

Problem 2

random variables:

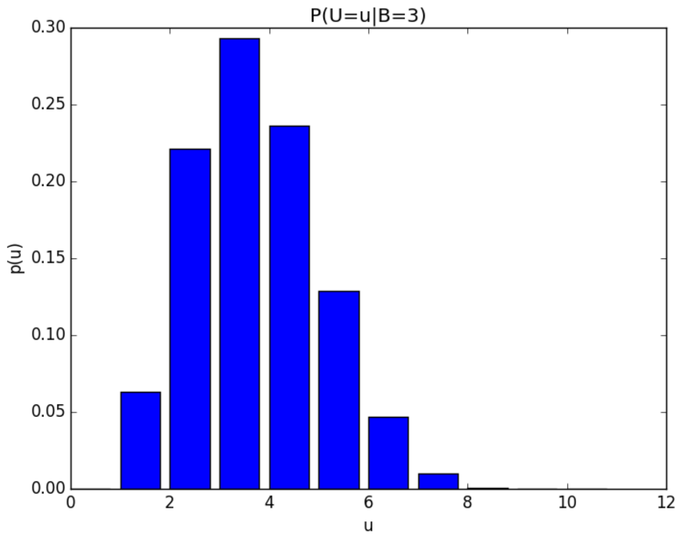
- U: the selected urn
- B: the number of black balls drawn from an urn

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

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

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

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



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

- $X=1$: next ball drawn from one of the urns is black
- $X=0$: next ball drawn from one of the urns is white

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

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

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

Problem 3

$$
P(X=x|\theta) = \theta^x(1-\theta)^{1-x}
$$

goal: determine θ_{MLE}

$$
\begin{aligned}
\theta_{MLE} &= \arg\max\nolimits_\theta P(X=x|\theta)\\
\Rightarrow 0 &= \frac{\mathrm{d}P}{\mathrm{d}\theta}(\theta_{MLE})\\
\Rightarrow 0 &= \frac{\mathrm{d}\log P}{\mathrm{d}\theta}(\theta_{MLE})
\end{aligned}
$$

$$
\begin{aligned}
\log P(X=x|\theta) &= \log\theta^x(1-\theta)^{1-x}\\
&= x\cdot\log\theta + (1-x)\cdot\log(1-\theta)
\end{aligned}
$$

$$
\begin{aligned}
0 &= \frac{\mathrm{d}}{\mathrm{d}\theta}(x\cdot\log\theta + (1-x)\cdot\log(1-\theta))\\
&= x\cdot\frac{1}{\theta} + (1-x)\cdot\frac{1}{1-\theta}\cdot(-1)\\
&= \frac{x}{\theta} - \frac{1-x}{1-\theta}\\
\Leftrightarrow \frac{1-x}{1-\theta} &= \frac{x}{\theta}\\
\Leftrightarrow \frac{1-x}{x} &= \frac{1-\theta}{\theta}\\
\Leftrightarrow \frac{1}{x} - 1 &= \frac{1}{\theta} - 1\\
\Leftrightarrow \frac{1}{x} &= \frac{1}{\theta}\\
\Leftrightarrow \theta &= x\\
\Rightarrow \theta_{MLE} &= x
\end{aligned}
$$

Problem 4

Show:

$$\begin{aligned}\theta_{MLE} &= \arg\max\nolimits_\theta P(X=x|\theta)\\ \Rightarrow 0 &= \frac{\mathrm{d}P}{\mathrm{d}\theta}(\theta_{MLE})\\ \Rightarrow 0 &= \frac{\mathrm{d}\log P}{\mathrm{d}\theta}(\theta_{MLE})\end{aligned}$$

$$\begin{aligned}\log P(X=m|\theta) &= \log \binom{N}{m}\theta^m(1-\theta)^{n-m}\\ &= \log\binom{N}{m} + m\cdot\log\theta + (N-m)\cdot\log(1-\theta)\end{aligned}$$

$$\begin{aligned}0 &= \frac{\mathrm{d}}{\mathrm{d}\theta}(\log\binom{N}{m} + m\cdot\log\theta + (N-m)\cdot\log(1-\theta))\\ &= m\cdot\frac{1}{\theta} + (N-m)\cdot\frac{1}{1-\theta}\cdot(-1)\\ &= \frac{m}{\theta} - \frac{N-m}{1-\theta}\\ \Leftrightarrow \frac{N-m}{1-\theta} &= \frac{m}{\theta}\\ \Leftrightarrow \frac{N-m}{m} &= \frac{1-\theta}{\theta}\\ \Leftrightarrow \frac{N-m}{m} &= \frac{1}{\theta}-1\\ \Leftrightarrow \frac{N-m}{m}+1 &= \frac{1}{\theta}\\ \Leftrightarrow \frac{N-m+m}{m} &= \frac{1}{\theta}\\ \Leftrightarrow \theta &= \frac{m}{N}\\ \Rightarrow \theta_{MLE} &= \frac{|X=1|}{|X=1|+|X=0|}\end{aligned}$$

Formalized Problem:
$mean(P(\theta|D)) = \lambda * mean(p(\theta)) + (1-\lambda) * \theta_{MLH}$
$mean(Beta(a+|X=1|, b+|X=0|)) = \lambda * mean(Beta(a,b)) + (1-\lambda) * \theta_{MLH}$
$\frac{a+|X=1|}{a+|X=1|+b+|X=0|} = \lambda * \frac{a}{a+b} + (1-\lambda) * \frac{|X=1|}{|X=1|+|X=0|}$

With: $0 \le \lambda \le 1$

Solution:

$$
\begin{aligned}
&\frac{a+|X=1|}{a+|X=1|+b+|X=0|} \\
&= \frac{a}{a+|X=1|+b+|X=0|} + \frac{|X=1|}{a+|X=1|+b+|X=0|} \\
&= \frac{a}{a+b+|X=1|+|X=0|} + \frac{|X=1|}{|X=1|+|X=0|+a+b} \\
&= \frac{a}{a+b*(1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|*(1+\frac{a+b}{|X=1|+|X=0|}} \\
&= \frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * \frac{1}{(1+\frac{a+b}{|X=1|+|X=0|}} \\
&= \frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * \frac{|X=1|+|X=0|}{(|X=1|+|X=0|+(a+b)} \\
&= \frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * (\frac{|X=1|+|X=0|+a+b}{(|X=1|+|X=0|)+(a+b)}) \\
&- \frac{a+b}{(|X=1|+|X=0|+(a+b))} \\
&= \frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * (1-\frac{a+b}{(|X=1|+|X=0|+(a+b))}) \\
&= \frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * (1-\frac{1}{(\frac{|X=1|+|X=0|}{a+b}+1)})
\end{aligned}
$$

Insert:

$\lambda = \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}}$
$mean(P(\theta|D)) = \lambda * \frac{a}{a+b} + (1-\lambda) * \frac{|X=1|}{|X=1|+|X=0|}$

qed.

Problem 5

$P(\lambda|x) = \frac{P(X|\lambda)*P(\lambda)}{P(X)}$
$P(\lambda|x) = \frac{P(X|\lambda)*P(\lambda)}{\int_0^\infty (P(X|\lambda)*P(\lambda))d\lambda}$

Step 1:

$\int_0^\infty P(\lambda) d\lambda = 1$
$\int_0^\infty Gamma(a,b) d\lambda = 1$
$\int_0^\infty \frac{b^a}{\Gamma(a)} * \lambda^{a-1} * e^{-b\lambda} d\lambda = 1$
$\frac{b^a}{\Gamma(a)} * \int_0^\infty \lambda^{a-1} * e^{-b\lambda} d\lambda = 1$
$\int_0^\infty \lambda^{a-1} * e^{-b\lambda} d\lambda = \frac{\Gamma(a)}{b^a}$

Step2:

$$P(X) = \int_0^\infty (P(X|\lambda) * P(\lambda)) d\lambda$$
$\int_0^\infty (P(X|\lambda) * Gamma(a,b)) d\lambda =$
$\int_0^\infty \frac{e^{-\lambda} * \lambda^x}{x!} * \frac{b^a}{\Gamma(a)} * \lambda^{a-1} * e^{-b\lambda}) d\lambda =$
$\frac{b^a}{\Gamma(a) * x!} \int_0^\infty e^{-\lambda} * \lambda^x * \lambda^{a-1} * e^{-b\lambda}) d\lambda =$

$$\frac{b^a}{\Gamma(a) * x!} \int_0^\infty \lambda^{x+a-1} * e^{-(b+1)\lambda}) d\lambda =$$

Applying Step1:
$\frac{b^a}{\Gamma(a) * x!} * \frac{\Gamma(x+a)}{(b+1)^{x+a}}$

Putting it together:
$P(\lambda|x) = \frac{P(X|\lambda) * P(\lambda)}{P(X)} =$
$\frac{\frac{e^{-\lambda} * \lambda^x}{x!} * \frac{b^a}{\Gamma(a)} * \lambda^{a-1} * e^{-b\lambda}}{\frac{b^a}{\Gamma(a) * x!} * \frac{\Gamma(x+a)}{(b+1)^{x+a}}} =$

$$\frac{\frac{b^a}{\Gamma(a) * x!} * e^{-\lambda} * \lambda^x * \lambda^{a-1} * e^{-b\lambda}}{\frac{b^a}{\Gamma(a) * x!} * \frac{\Gamma(x+a)}{(b+1)^{x+a}}}$$

$$\frac{e^{-\lambda} * \lambda^x * \lambda^{a-1} * e^{-b\lambda}}{\frac{\Gamma(x+a)}{(b+1)^{x+a}}}$$

$$e^{-\lambda} * \lambda^x * \lambda^{a-1} * e^{-b\lambda} * \frac{(b+1)^{x+a}}{\Gamma(x+a)} =$$

$\lambda^{a+x-1} * e^{-(b+1)\lambda} * \frac{(b+1)^{x+a}}{\Gamma(x+a)} =$
$Gamma(a+x, b+1)$
Calculating Map:
$\frac{d}{d\lambda} P(\lambda|x) = 0$
$\frac{d}{d\lambda} ln(P(\lambda|x)) = 0$
$\frac{d}{d\lambda} ln(P(\lambda|x)) = \frac{d}{d\lambda} ln(\frac{(b+1)^{x+a}}{\Gamma(x+a)} * \lambda^{a+x-1} * e^{-(b+1)\lambda}) =$
$\frac{d}{d\lambda} ln(\frac{(b+1)^{x+a}}{\Gamma(x+a)}) + ln(\lambda^{a+x-1}) + ln(e^{-(b+1)\lambda}) =$
$\frac{d}{d\lambda}(a+x-1) * ln(\lambda) + -(b+1)\lambda =$
$(a+x-1) * \frac{1}{\lambda} + -(b+1)$

$(a+x-1) * \frac{1}{\lambda} + -(b+1) = 0$
$(a+x-1) * \frac{1}{\lambda} = (b+1)$
$(a+x-1) = (b+1)\lambda$

$\lambda = \frac{a+x+1}{b+1}$
$\Theta_{MAP} = \frac{a+x+1}{b+1}$