Machine Learning Worksheet 4

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

November 8, 2015

Problem 1

- T: number of tails thrown before the first occurrance of a head
- coin is unbiased $\Rightarrow$ probability of throwing head or tail in a single coin flip is $\frac{1}{2}$
- $\mathrm{E}[H] = 1$ since the expriment is stopped after the first head

$$\begin{aligned} P(T = t) &= \left(\frac{1}{2}\right)^t \cdot \frac{1}{2} \\ &= \left(\frac{1}{2}\right)^{t+1} \end{aligned}$$

$$\begin{aligned} \mathrm{E}[T] &= \int_0^\infty P(T = t) \cdot t \, \mathrm{d}t \\ &= \int_0^\infty \left(\frac{1}{2}\right)^{t+1} \cdot t \, \mathrm{d}t \\ &= \frac{1}{2} \cdot \int_0^\infty \left(\frac{1}{2}\right)^t \cdot t \, \mathrm{d}t \\ &= \frac{1}{2} \cdot \int_0^\infty e^{\ln\left(\frac{1}{2}\right) \cdot t} \cdot t \, \mathrm{d}t \end{aligned}$$

Apply the product rule of integration:

$$\begin{aligned}
\mathrm{E}[T] &= \frac{1}{2} \cdot \int_0^\infty e^{\ln\left(\frac{1}{2}\right)\cdot t} \cdot t \,\mathrm{d}t \\
&= \frac{1}{2} \cdot \left(\left[\frac{1}{\ln\left(\frac{1}{2}\right)} \cdot e^{\ln\left(\frac{1}{2}\right)\cdot t} \cdot t \right]_{t=0}^\infty - \int_0^\infty \frac{1}{\ln\left(\frac{1}{2}\right)} \cdot e^{\ln\left(\frac{1}{2}\right)\cdot t} \cdot 1 \,\mathrm{d}t \right) \\
&= \frac{1}{2} \cdot \left((0-0) - \frac{1}{\ln\left(\frac{1}{2}\right)} \cdot \int_0^\infty e^{\ln\left(\frac{1}{2}\right)\cdot t} \,\mathrm{d}t \right) \\
&= -\frac{1}{2} \cdot \left[\frac{1}{\left(\ln\left(\frac{1}{2}\right)\right)^2} \cdot e^{\ln\left(\frac{1}{2}\right)\cdot t} \right]_0^\infty \\
&= -\frac{1}{2} \cdot \frac{1}{\left(\ln\left(\frac{1}{2}\right)\right)^2} \cdot (0 - e^0) \\
&= \frac{1}{2 \cdot \left(\ln\left(\frac{1}{2}\right)\right)^2} \\
&\approx 1.04068
\end{aligned}$$

Problem 2

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

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

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

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

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

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

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

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

Problem 3

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

goal: determine θ_{MLE} for n experiments
multiple bernoulli experiments $\Rightarrow$ binomial distribution

- Z: number of experiments that end with $X=1$

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

$$
\begin{aligned}
P(Z=z|\theta) &= \binom{n}{z}\theta^z(1-\theta)^{n-z} \\
\log P(Z=z|\theta) &= \log\left(\binom{n}{z}\theta^z(1-\theta)^{n-z}\right) \\
&= \log\binom{n}{z} + z\log\theta + (n-z)\log(1-\theta)
\end{aligned}
$$

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

Problem 4

Show: *TODO: Show $\theta_{MLH} = \frac{|X=1|}{|X=1|+|X=0|}$ referece to prob 3

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: $0 \leq \lambda \leq 1$

Solution: $\frac{a+|X=1|}{a+|X=1|+b+|X=0|} =$
$\frac{a}{a+|X=1|+b+|X=0|} + \frac{|X=1|}{a+|X=1|+b+|X=0|} =$
$\frac{a}{a+b+|X=1|+|X=0|} + \frac{|X=1|}{|X=1|+|X=0|+a+b} =$
$\frac{a}{a+b*(1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|*(1+\frac{a+b}{|X=1|+|X=0|}} =$
$\frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * \frac{1}{(1+\frac{a+b}{|X=1|+|X=0|}} =$
$\frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * \frac{|X=1|+|X=0|}{(|X=1|+|X=0|+(a+b)} =$
$\frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * (\frac{|X=1|+|X=0|+a+b}{(|X=1|+|X=0|+(a+b)} - \frac{a+b}{(|X=1|+|X=0|+(a+b))} =$
$\frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * (1 - \frac{a+b}{(|X=1|+|X=0|+(a+b))}) =$

$\frac{a}{a+b} * \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}} + \frac{|X=1|}{|X=1|+|X=0|} * (1 - \frac{1}{(\frac{|X=1|+|X=0|}{a+b}+1)}) =$
insert:
$\lambda = \frac{1}{1+\frac{|X=1|+|X=0|}{a+b}}$
$= \lambda * \frac{a}{a+b} + (1-\lambda) * \frac{|X=1|}{|X=1|+|X=0|}$

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:
$\frac{d}{d\lambda} P(\lambda|x) = 0$

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

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