Neuroinformatics Exercise
Fall 2016

Lecture E6

Name:
Olivera Stojanović, Pascal Nieters

Instructions: (20 points) Solve the following problems. Write clearly and use same symbols as used in the lecture. Add comments, explanations or questions to your solution if necessary. Add additional paper sheets if space on these exercise sheets is insufficient. Label the additional sheets with your name and matriculation number. Each exercise sheet has in total 20 standard credit points. However, on some sheets you might find problems that can give you extra points. Such a problem is marked with ADVANCED. At the end of the term the total sum of standard and extra points needs to be larger than 60% of all standard points to get admitted to the exam at the end of the term. The deadline for this exercise sheet is: **16.12.2016**
Please, staple your sheets!

(8pts) **1.** Let x be a discrete random variable taking on the values $k = 0, 1, 2, \ldots$ Suppose x has a Poisson distribution with rate parameter $\lambda > 0$, such that $p(x = k) = p(k|\lambda) = \frac{\lambda^k \exp(-\lambda)}{k!}$. Show that $\mathbb{E}[x] = \lambda$.
Hint: use the equivalences $\sum_{x=0}^{\infty} \frac{\lambda^x}{x!} = \exp(\lambda)$ and $0! = 1$.

8 pts

Solution:

$$
\begin{aligned}
\mathbb{E}[x] &= \sum_{x=0}^{\infty} x \cdot p(x|\lambda) \\
&= \sum_{x=0}^{\infty} x \cdot \frac{\lambda^x \exp(-\lambda)}{x!} \\
&= 0 + \sum_{x=1}^{\infty} x \cdot \frac{\lambda^x \exp(-\lambda)}{x!} \\
&= \sum_{x=1}^{\infty} \frac{\lambda^x \exp(-\lambda)}{(x-1)!} \\
&= \exp(-\lambda) \sum_{x=1}^{\infty} \frac{\lambda^x}{(x-1)!} \\
&= \lambda \exp(-\lambda) \sum_{x=1}^{\infty} \frac{\lambda^{x-1}}{(x-1)!} \\
&= \lambda \exp(-\lambda) \underbrace{\sum_{x=0}^{\infty} \frac{\lambda^x}{x!}}_{\exp(\lambda)} \\
&= \lambda
\end{aligned}
$$

(6pts) **2.** We have a single spiketrain with spikes described by a binary process. Use a Bernoulli model with the assumption of i.i.d. bins to calculate the ML estimate of the probability of a spike per bin.

6 pts

1	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	1	0	0	0

Solution:

$$
\begin{aligned}
p(x|\mu) &= \prod_{i=1}^{N} \mu^{x_i}(1-\mu)^{1-x_i} && \text{with } x_i \in \{0,1\},\ N \text{ random variables} \\
\ln p(x|\mu) &\overset{\star}{=} \ln\left(\prod_{i=1}^{N} \mu^{x_i}(1-\mu)^{1-x_i}\right) \\
&= \sum_{i=1}^{N} \ln\left(\mu^{x_i}(1-\mu)^{1-x_i}\right) \\
&= \sum_{i=1}^{N} \left(\ln \mu^{x_i} + \ln(1-\mu)^{1-x_i}\right) \\
&= \sum_{i=1}^{N} \left(\underbrace{x_i \ln \mu}_{\log a^x = x \log a} + (1-x_i)\ln(1-\mu)\right)
\end{aligned}
$$

For $\star$ use the log trick.

First, we derive the partial derivative with respect to μ.

$$\frac{\partial}{\partial\mu} \ln p(x|\mu) = \sum_{i=1}^{N} \left(x_i \frac{\partial}{\partial\mu} \ln \mu + (1-x_i)\frac{\partial}{\partial\mu}\ln(1-\mu)\right) \star\star$$

Trick for $\star\star$: use the chain rule.
Substitution $f(y(\mu))$:

$$
\begin{aligned}
y(\mu) &= 1-\mu \\
f(y) &= \ln(y(\mu)) \\
&= \ln(1-\mu) \\
\frac{d}{d\mu} \cdot f(y(\mu)) &= \frac{df}{dy} \cdot \frac{dy}{d\mu} \\
\frac{df}{dy} &= \frac{1}{y} = \frac{1}{1-\mu} \\
\frac{dy}{d\mu} &= -1 \\
\frac{df}{dy} \cdot \frac{dy}{d\mu} &= \frac{1}{1-\mu} \cdot (-1)
\end{aligned}
$$

We obtain the first derivative:

$$\frac{\partial}{\partial\mu} \ln p(x|\mu) = \sum_{i=1}^{N} \left(x_i \frac{1}{\mu} + (1-x_i)\frac{-1}{(1-\mu)}\right)$$

Now we set this partial derivative to zero:

$$
\begin{aligned}
0 &= \sum_{i=1}^{N} \left(x_i \frac{1}{\mu_{ML}} + (1 - x_i) \frac{-1}{(1 - \mu_{ML})} \right) \\
&\overset{\star\star\star}{=} \sum_{i=1}^{N} \left(x_i (1 - \mu_{ML}) - \mu_{ML}(1 - x_i) \right) \\
&= \sum_{i=1}^{N} \left(x_i - x_i \mu_{ML} - \mu_{ML} + \mu_{ML} x_i \right) \\
&= \sum_{i=1}^{N} \left(x_i - \mu_{ML} \right) \\
&= -N\mu_{ML} - \underbrace{\sum_{i=1}^{N} x_i}_{m} \\
\mu_{ML} &= \frac{m}{N}
\end{aligned}
$$

Trick for $\star\star\star$: multiply both sides of equation with $\mu_{ML}(1 - \mu_{ML})$.

(6[pts]) **3.** Derive the observed Fisher information for the Bernoulli, the Binomial and the Poisson distribution. *Solution*:

6 pts

1. The Bernoulli distribution:

$$
p(x|\mu) = \prod_{i=1}^{N} \mu^{x_i} (1 - \mu)^{1 - x_i} \qquad \text{with } x_i \in \{0, 1\},\ N \text{ random variables}
$$

The log-likelihood function of the Bernoulli distribution:

$$
\begin{aligned}
\ln p(x|\mu) &= \ln \left(\prod_{i=1}^{N} \mu^{x_i} (1 - \mu)^{1 - x_i} \right) \\
&= \sum_{i=1}^{N} \left(x_i \ln \mu + (1 - x_i) \ln(1 - \mu) \right)
\end{aligned}
$$

(Observed) Fisher information for the likelihood of Bernoulli distribution samples in N trials:

$$\begin{aligned}
I(\mu) &= -\frac{\partial^2}{\partial^2\mu}\ln p(x|\mu) \\
&= -\frac{\partial^2}{\partial^2\mu}\left(\sum_{i=1}^{N}(x_i\ln\mu + (1-x_i)\ln(1-\mu))\right) \\
&= -\frac{\partial}{\partial\mu}\left(\frac{\partial}{\partial\mu}\left(\sum_{i=1}^{N}(x_i\ln\mu + (1-x_i)\ln(1-\mu))\right)\right) \\
&= -\frac{\partial}{\partial\mu}\left(\sum_{i=1}^{N}\left(x_i\frac{1}{\mu} + (1-x_i)\frac{-1}{(1-\mu)}\right)\right) \\
&= -\sum_{i=1}^{N}\left(-\frac{x_i}{\mu^2} - \frac{(1-x_i)}{(1-\mu)^2}\right)
\end{aligned}$$

We set the fisrt derivative (score function) to zero to get MLE:

$$\begin{aligned}
0 &= \sum_{i=1}^{N}\left(x_i\frac{1}{\mu_{ML}} + (1-x_i)\frac{-1}{(1-\mu_{ML})}\right) \\
&\overset{\star}{=} \sum_{i=1}^{N}(x_i(1-\mu_{ML}) - \mu_{ML}(1-x_i)) \\
&= \sum_{i=1}^{N}(x_i - x_i\mu_{ML} - \mu_{ML} + \mu_{ML}x_i) \\
&= \sum_{i=1}^{N}(x_i - \mu_{ML}) \\
&= -N\mu_{ML} - \underbrace{\sum_{i=1}^{N}x_i}_{m} \\
\mu_{ML} &= \frac{m}{N}
\end{aligned}$$

Trick for $\star$: multiply both sides of equation with $\mu_{ML}(1-\mu_{ML})$.

Then we plug into the solution for the Fisher information:

$$
\begin{aligned}
I(\mu_{ML}) &= -\sum_{i=1}^{N}\left(-\frac{x_i}{\mu_{ML}^2} - \frac{(1-x_i)}{(1-\mu_{ML})^2}\right) \\
&= \frac{\sum_{i=1}^{N} x_i}{\mu_{ML}^2} + \frac{(N - \sum_{i=1}^{N} x_i)}{(1-\mu_{ML})^2} \\
&= \frac{m}{\mu_{ML}^2} + \frac{(N-m)}{(1-\mu_{ML})^2} \\
&= \frac{m}{\frac{m^2}{N^2}} + \frac{(N-m)}{(1-\frac{m}{N})^2} \\
&= \frac{N^2}{m(1-\frac{m}{N})} \\
&= \frac{N}{\mu_{ML}(1-\mu_{ML})}
\end{aligned}
$$

2. The Binomial distribution:

$$
\begin{aligned}
m &= \sum_{i=1}^{N} x_i \;^{\star} \\
p(m|N,\mu) &= \binom{N}{m}\mu^m\,(1-\mu)^{(N-m)}
\end{aligned}
$$

$\star$ **Note:** in the lecture "our" m is called k.

The log-likelihood function of the Binomial distribution:

$$
\begin{aligned}
\ln p(m|N,\mu) &= \ln\left(\binom{N}{m}\mu^m\,(1-\mu)^{(N-m)}\right) \\
&= \ln\binom{N}{m} + m\ln\mu + (N-m)\ln(1-\mu)
\end{aligned}
$$

(Observed) Fisher information for the Binomial distribution:

$$
\begin{aligned}
I(\mu) &= -\frac{\partial^2}{\partial^2\mu}\ln p(m|\mu) \\
&= -\frac{\partial}{\partial\mu}\left(\frac{\partial}{\partial\mu}\left(\ln\binom{N}{m} + m\ln\mu + (N-m)\ln(1-\mu)\right)\right) \\
&= -\frac{\partial}{\partial\mu}\left(\frac{m}{\mu} - \frac{(N-m)}{1-\mu}\right) \\
&= \frac{m}{\mu^2} + \frac{(N-m)}{(1-\mu)^2}
\end{aligned}
$$

We set the first derivative (score function) to zero:

$$\begin{aligned}
0 &= \frac{m}{\mu_{ML}} - \frac{N-m}{1-\mu_{ML}} \\
&= m(1-\mu_{ML}) - \mu_{ML}(N-m) \\
&= m - \mu_{ML}m - \mu_{ML}N + \mu_{ML}m \\
\mu_{ML} &= \frac{m}{N}
\end{aligned}$$

Plug into the solution for the Fisher information:

$$\begin{aligned}
I(\mu_{ML}) &= \frac{m}{\mu_{ML}^2} + \frac{(N-m)}{(1-\mu_{ML})^2} \\
&= \frac{m}{\frac{m^2}{N^2}} + \frac{(N-m)}{(1-\frac{m}{N})^2} \\
&= \frac{N^2}{m} + \frac{(N-m)}{(1-\frac{m}{N})^2} \\
&= \frac{N^2}{m(1-\frac{m}{N})} \\
&= \frac{N}{\mu_{ML}(1-\mu_{ML})}
\end{aligned}$$

3. The Poisson distribution:

$$p(m|\lambda) = \frac{\lambda^m \exp(-\lambda)}{m!}$$

The log-likelihood function of the Poisson distribution:

$$\begin{aligned}
\ln p(m|\lambda) &= \ln\left(\exp(-\lambda)\frac{\lambda^m}{m!}\right) \\
&= -\lambda + m\ln\lambda - \ln m!
\end{aligned}$$

(Observed) Fisher information for the Poisson distribution:

$$\begin{aligned}
I(\lambda) &= -\frac{\partial^2}{\partial^2\lambda}\ln p(m|\lambda) \\
&= -\frac{\partial}{\partial\lambda}\left(\frac{\partial}{\partial\lambda}\left(-\lambda + m\ln\lambda - \ln m!\right)\right) \\
&= -\frac{\partial}{\partial\lambda}\left(-1 + \frac{m}{\lambda}\right) \\
&= \frac{m}{\lambda^2}
\end{aligned}$$

Then we set the first derivative (score function) to zero:

$$\begin{aligned}
0 &= -1 + \frac{m}{\lambda_{ML}} \\
\lambda_{ML} &= m
\end{aligned}$$

Plug into the solution for the Fisher information:

$$\begin{aligned} I(\lambda_{ML}) &= \frac{m}{\lambda_{ML}^2} \\ &= \frac{1}{\lambda_{ML}} \end{aligned}$$