Solutions for Problem Sheet 4 (2014 Aachen Graduierten-Kolleg)

1(a) [10 marks] The exponentially distributed time measurements, $t_1, \ldots, t_n$, and the Gaussian distributed calibration measurement y are all independent, so the likelihood is simply the product of the corresponding pdfs:

$$L(\tau, \lambda) = \prod_{i=1}^{n} \frac{1}{\tau + \lambda} e^{-t_i/(\tau+\lambda)} \frac{1}{\sqrt{2\pi}\sigma} e^{-(y-\lambda)^2/2\sigma^2} \; .$$

The log-likelihood is therefore

$$\ln L(\tau, \lambda) = -n \ln(\tau + \lambda) - \frac{1}{\tau + \lambda} \sum_{i=1}^{n} t_i - \frac{(y-\lambda)^2}{2\sigma^2} + C \; ,$$

where C represents terms that do not depend on the parameters and therefore can be dropped. Differentiating $\ln L$ with respect to the parameters gives

$$\begin{aligned} \frac{\partial \ln L}{\partial \tau} &= -\frac{n}{\tau + \lambda} + \frac{\sum_{i=1}^{n} t_i}{(\tau + \lambda)^2} \\ \frac{\partial \ln L}{\partial \lambda} &= -\frac{n}{\tau + \lambda} + \frac{\sum_{i=1}^{n} t_i}{(\tau + \lambda)^2} + \frac{y - \lambda}{\sigma^2} \; . \end{aligned}$$

Setting the derivatives to zero and solving for τ and λ gives the ML estimators,

$$\begin{aligned} \hat{\tau} &= \frac{1}{n} \sum_{i=1}^{n} t_i - y \\ \hat{\lambda} &= y \; . \end{aligned}$$

1(b) [10 marks] The variances of $\hat{\lambda}$ and $\hat{\tau}$ and their covariance are

$$\begin{aligned} V[\hat{\lambda}] &= V[y] = \sigma^2 \; , \\ V[\hat{\tau}] &= V\left[\frac{1}{n} \sum_{i=1}^{n} t_i - y\right] = \frac{1}{n^2} \sum_{i=1}^{n} V[t_i] + V[y] = \frac{(\tau + \lambda)^2}{n} + \sigma^2 \\ \mathrm{cov}[\hat{\tau}, \hat{\lambda}] &= \mathrm{cov}\left[\frac{1}{n} \sum_{i=1}^{n} t_i - y, y\right] = -V[y] = -\sigma^2 \; , \end{aligned}$$

For the covariance we used the fact that t_i and y are independent and thus have zero covariance.

1(c) [10 marks] The standard deviations of $\hat{\tau}$ and $\hat{\lambda}$ can be determined from the contour of $\ln L(\tau, \lambda) = \ln L_{\max} - 1/2$, as shown in Fig. 1. The standard can be approximated by the distance from the maximum of $\ln L$ to the tangent line to the contour (in either direction).

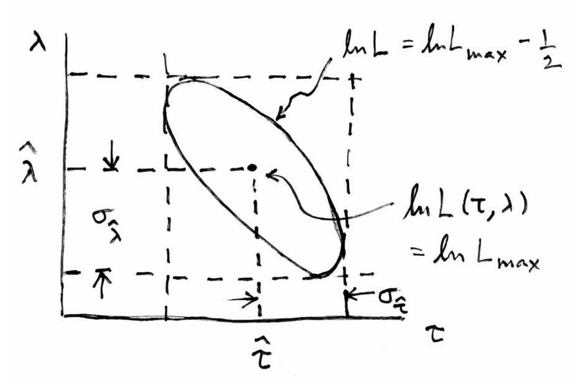

Figure 1: Illustration of the method to find $\sigma_{\hat{\tau}}$ and $\sigma_{\hat{\lambda}}$ from the contour of $\ln L(\tau, \lambda) = \ln L_{\max} - 1/2$ (see text).

If λ were to be known exactly, then the standard deviation of $\hat{\tau}$ would be less. This can be seen from Fig. 1, for example, since the distance one need to move τ away from the maximum of $\ln L$ to get to $\ln L_{\max} - 1/2$ would be less if λ were to be fixed at $\hat{\lambda}$.

1(d) [10 marks] The second derivatives of lnL are

$$\begin{aligned}
\frac{\partial^2 \ln L}{\partial \tau^2} &= \frac{n}{(\tau+\lambda)^2} - \frac{2\sum_{i=1}^n t_i}{(\tau+\lambda)^3} \,, \\
\frac{\partial^2 \ln L}{\partial \lambda^2} &= \frac{n}{(\tau+\lambda)^2} - \frac{2\sum_{i=1}^n t_i}{(\tau+\lambda)^3} - \frac{1}{\sigma^2} \,, \\
\frac{\partial^2 \ln L}{\partial \tau \partial \lambda} &= \frac{n}{(\tau+\lambda)^2} - \frac{2\sum_{i=1}^n t_i}{(\tau+\lambda)^3} \,.
\end{aligned}$$

Using $E[t_i] = \tau + \lambda$ we find the expectation values of the second derivatives,

$$\begin{aligned}
E\left[\frac{\partial^2 \ln L}{\partial \tau^2}\right] &= \frac{n}{(\tau+\lambda)^2} - \frac{2n(\tau+\lambda)}{(\tau+\lambda)^3} = -\frac{n}{(\tau+\lambda)^2} \,, \\
E\left[\frac{\partial^2 \ln L}{\partial \lambda^2}\right] &= -\frac{n}{(\tau+\lambda)^2} - \frac{1}{\sigma^2} \,, \\
E\left[\frac{\partial^2 \ln L}{\partial \tau \partial \lambda}\right] &= -\frac{n}{(\tau+\lambda)^2} \,.
\end{aligned}$$

The inverse covariance matrix of the estimators is given by

$$V_{ij}^{-1} = -E\left[\frac{\partial^2 \ln L}{\partial \theta_i \partial \theta_j}\right]$$

where here we can take, e.g., $\theta_1 = \tau$ and $\theta_2 = \lambda$. We are given the formula for the inverse of the corresponding 2×2 matrix, and by substituting in the ingredients we find

$$V = \begin{pmatrix} \frac{(\tau+\lambda)^2}{n} + \sigma^2 & -\sigma^2 \\ -\sigma^2 & \sigma^2 \end{pmatrix}$$

which are the same as what was found in (c).

2(a) [10 marks] The likelihood function in terms of $\nu_{\rm a}$ and $\nu_{\rm b}$ is the product of two Poisson terms,

$$L(\nu_{\rm a}, \nu_{\rm b}) = \frac{\nu_{\rm a}^{n_{\rm a}}}{n_{\rm a}!} e^{-\nu_{\rm a}} \frac{\nu_{\rm b}^{n_{\rm b}}}{n_{\rm b}!} e^{-\nu_{\rm b}} \ .$$

The log-likelihood is therefore

$$\ln L(\nu_{\rm a}, \nu_{\rm b}) = n_{\rm a} \ln \nu_{\rm a} - \nu_{\rm a} + n_{\rm b} \ln \nu_{\rm b} - \nu_{\rm b} + C \ ,$$

where C represents terms that do not depend on the parameters and thus can be dropped. The parameters $\nu_{\rm a}$ and $\nu_{\rm b}$ can be written in terms of ν and α as

$$\begin{aligned}
\nu_{\rm a} &= \frac{\nu}{2}(1+\alpha) \\
\nu_{\rm b} &= \frac{\nu}{2}(1-\alpha) \ ,
\end{aligned}$$

so that the log-likelihood is (dropping the constant C),

$$\begin{aligned}
\ln L(\nu, \alpha) &= n_{\rm a} \ln \left[\frac{\nu}{2}(1+\alpha) \right] - \frac{\nu}{2}(1+\alpha) + n_{\rm b} \ln \left[\frac{\nu}{2}(1-\alpha) \right] - \frac{\nu}{2}(1-\alpha) \\
&= (n_{\rm a} + n_{\rm b}) \ln \nu - \nu + n_{\rm a} \ln(1+\alpha) + n_{\rm b} \ln(1-\alpha) \ .
\end{aligned}$$

The derivatives with respect to ν and α are

$$\begin{aligned}
\frac{\partial \ln L}{\partial \nu} &= \frac{n_{\rm a} + n_{\rm b}}{\nu} - 1 \ , \\
\frac{\partial \ln L}{\partial \alpha} &= \frac{n_{\rm a}}{1+\alpha} - \frac{n_{\rm b}}{1-\alpha} \ .
\end{aligned}$$

Setting the derivatives to zero and solving for ν and α gives the ML estimators,

$$\begin{aligned}
\hat{\nu} &= n_{\rm a} + n_{\rm b} \\
\hat{\alpha} &= \frac{n_{\rm a} - n_{\rm b}}{n_{\rm a} + n_{\rm b}} \ .
\end{aligned}$$

2(b) [10 marks] Using error propagation, the variance of $\hat{\alpha}$ can be approximated as

$$V[\hat{\alpha}] \approx \left(\frac{\partial \hat{\alpha}}{\partial n_{\rm a}} \right)^2 \Bigg|_{\mathbf{n}=\boldsymbol{\nu}} V[n_{\rm a}] + \left(\frac{\partial \hat{\alpha}}{\partial n_{\rm b}} \right)^2 \Bigg|_{\mathbf{n}=\boldsymbol{\nu}} V[n_{\rm b}] \ ,$$

Computing the derivatives, which are evaluated at $n_{\rm a} = \nu_{\rm a}$ and $n_{\rm b} = \nu_{\rm b}$, and using $V[n_{\rm a}] = \nu_{\rm a}$ and $V[n_{\rm b}] = \nu_{\rm b}$ gives

$$
\begin{aligned}
V[\hat{\alpha}] &= \left(\frac{2\nu_{\rm b}}{\nu^2}\right)^2 \nu_{\rm a} + \left(\frac{2\nu_{\rm a}}{\nu^2}\right)^2 \nu_{\rm b} \\
&= \left(\frac{\nu(1-\alpha)}{\nu^2}\right)^2 \frac{\nu}{2}(1+\alpha) + \left(\frac{\nu(1+\alpha)}{\nu^2}\right)^2 \frac{\nu}{2}(1-\alpha) \\
&= \frac{1-\alpha^2}{\nu} \; .
\end{aligned}
$$

2(c) [14 marks] Writing the likelihood in terms of ν and α (see, e.g., $\ln L$ from (a)) gives

$$L(\nu, \alpha) \propto \nu^{(n_{\rm a}+n_{\rm b})} e^{-\nu} (1+\alpha)^{n_{\rm a}} (1-\alpha)^{n_{\rm b}} \; .$$

Using the prior given, $\pi(\nu, \alpha) \propto 1/\sqrt{\nu}$, the joint posterior for α and ν is

$$p(\nu, \alpha | n_{\rm a}, n_{\rm b}) \propto \nu^{(n_{\rm a}+n_{\rm b}-1/2)} e^{-\nu} (1+\alpha)^{n_{\rm a}} (1-\alpha)^{n_{\rm b}} \; .$$

This factorizes into a function of ν times a function of α, so we can therefore conclude α and ν are independent with

$$
\begin{aligned}
p(\nu, \alpha | n_{\rm a}, n_{\rm b}) &= p(\nu | n_{\rm a}, n_{\rm b}) p(\alpha | n_{\rm a}, n_{\rm b}) \; , \\
p(\nu | n_{\rm a}, n_{\rm b}) &\propto \nu^{(n_{\rm a}+n_{\rm b}-1/2)} e^{-\nu} \; , \\
p(\alpha | n_{\rm a}, n_{\rm b}) &\propto (1+\alpha)^{n_{\rm a}} (1-\alpha)^{n_{\rm b}} \; .
\end{aligned}
$$

2(d) [6 marks] The posterior modes for α and ν are found by setting the corresponding derivatives to zero:

$$
\begin{aligned}
\frac{\partial p(\nu | n_{\rm a}, n_{\rm b})}{\partial \nu} &\propto \left(n_{\rm a} + n_{\rm b} - \frac{1}{2}\right) \nu^{(n_{\rm a}+n_{\rm b}-3/2)} e^{-\nu} - \nu^{(n_{\rm a}+n_{\rm b}-1/2)} e^{-\nu} = 0 \; , \\
\frac{\partial p(\alpha | n_{\rm a}, n_{\rm b})}{\partial \alpha} &\propto (1+\alpha)^{n_{\rm a}} n_{\rm b} (1-\alpha)^{n_{\rm b}-1} (-1) + n_{\rm a} (1-\alpha)^{n_{\rm b}} (1+\alpha)^{n_{\rm a}-1} = 0 \; .
\end{aligned}
$$

Solving for ν and α gives the Bayesian highest probability density (HPD) estimators,

$$
\begin{aligned}
\hat{\nu}_{\rm Bayes} &= n_{\rm a} + n_{\rm b} - \frac{1}{2} \; , \\
\hat{\alpha}_{\rm Bayes} &= \frac{n_{\rm a} - n_{\rm b}}{n_{\rm a} + n_{\rm b}} \; .
\end{aligned}
$$

The posterior mode for α is the same as the ML estimator, which follows from the fact that the prior for α and ν was taken to be independent of α. As the prior does, however, depend on ν, one does not expect the ML estimator and posterior mode to agree for ν.