Statistical Methods for Particle Physics
Graduierten-Kolleg RWTH Aachen, February 2014
Problem sheet 5 solutions

1(a) [6 marks] The likelihood function is given by the binomial distribution evaluated with the single observed value n and regarded as a function of the unknown parameter θ:

$$L(\theta) = \frac{N!}{n!(N-n)!}\theta^n (1-\theta)^{N-n} .$$

The log-likelihood function is therefore (3 marks)

$$\ln L(\theta) = n \ln \theta + (N-n) \ln(1-\theta) + C ,$$

where C represents terms not depending on θ. Setting the derivative of $\ln L$ equal to zero,

$$\frac{\partial \ln L}{\partial \theta} = \frac{n}{\theta} - \frac{N-n}{1-\theta} = 0 ,$$

we find the ML estimator to be (3 marks)

$$\hat{\theta} = \frac{n}{N} .$$

1(b) [8 marks] We are given the expectation and variance of a binomial distributed variable as $E[n] = N\theta$ and $V[n] = N\theta(1-\theta)$. Using these results we find the expectation value of $\hat{\theta}$ to be (4 marks)

$$E[\hat{\theta}] = E\left[\frac{n}{N}\right] = \frac{E[n]}{N} = \frac{N\theta}{N} = \theta ,$$

and therefore the bias is $b = E[\hat{\theta}] - \theta = 0$. Similarly we find the variance to be

$$V[\hat{\theta}] = V\left[\frac{n}{N}\right] = \frac{1}{N^2}V[n] = \frac{N\theta(1-\theta)}{N^2} = \frac{\theta(1-\theta)}{N} .$$

1(c) [8 marks] Suppose we observe $n = 0$ for $N = 10$ trials. The upper limit on θ at a confidence level of $\text{CL} = 1 - \alpha$ is the value of θ for which there is a probability α to find as few events as we found or fewer, i.e., (6 marks)

$$\alpha = P(n \leq 0; N, \theta) = \frac{N!}{0!(N-0)!}\theta^0 (1-\theta)^{N-0} .$$

Solving for θ gives the 95% CL upper limit (2 marks)

$$\theta_{\text{up}} = 1 - \alpha^{1/N} = 1 - 0.05^{1/10} = 0.26 .$$

1(d) [10 marks] To find the Jeffreys prior we need the second derivative of $\ln L$, (2 marks)

$$\frac{\partial^2 \ln L}{\partial \theta^2} = -\frac{n}{\theta^2} - \frac{N-n}{(1-\theta)^2} \; .$$

The expected Fisher information is therefore (2 marks)

$$I(\theta) = -E\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right] = \frac{N\theta}{\theta^2} + \frac{N(1-\theta)}{(1-\theta)^2} = \frac{N}{\theta} + \frac{N}{1-\theta} = \frac{N}{\theta(1-\theta)} \; .$$

The Jeffreys prior is therefore (3 marks)

$$\pi(\theta) \propto \frac{1}{\sqrt{\theta(1-\theta)}} \; .$$

Using this in Bayes theorem to find the posterior pdf gives (3 marks)

$$p(\theta|n) \propto L(n|\theta)\pi(\theta) \propto \frac{\theta^n (1-\theta)^{N-n}}{\sqrt{\theta(1-\theta)}} = \theta^{n-1/2}(1-\theta)^{N-n-1/2} \; .$$

1(e) [8 marks] To find a Bayesian upper limit on θ one simply integrates the posterior pdf so that a specified probability $1-\alpha$ is contained below $\theta_{\rm up}$, i.e.,

$$1 - \alpha = \int_0^{\theta_{\rm up}} p(\theta|n)\, d\theta \; ,$$

solving for $\theta_{\rm up}$ gives the upper limit. (3 marks)

A frequentist upper limit as found in (c) is a function of the data designed to be greater than the true value of the parameter with a fixed probability (the confidence level) regardless of the parameter's actual value. A Bayesian interval can be regarded as reflecting a range for the parameter where it is believed to lie with a fixed probability (the credibility level). Note that with the Jeffreys prior, one may not necessary use the degree of belief interpretation of the interval, but rather take it to have a certain probability to cover the true θ (which in general will depend on θ). (5 marks)

2(a) [4 points] The likelihood function is

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta) = \prod_{i=1}^{n} \frac{x_i}{\theta^2} e^{-x_i/\theta} \ .$$

Taking the logarithm gives

$$\ln L(\theta) = \sum_{i=1}^{n} \left(\ln \frac{x_i}{\theta^2} - \frac{x_i}{\theta} \right) = \sum_{i=1}^{n} \left(-2\ln\theta + \ln x_i - \frac{x_i}{\theta} \right) \ .$$

To find the ML estimator we set the derivative of $\ln L$ with respect to θ equal to zero,

$$\frac{\partial}{\partial\theta} \ln L(\theta) = -\frac{2n}{\theta} + \frac{1}{\theta^2} \sum_{i=1}^{n} x_i = 0 \ .$$

Solving for θ gives the ML estimator,

$$\hat{\theta} = \frac{1}{2n} \sum_{i=1}^{n} x_i = \frac{\overline{x}}{2} \ .$$

2(b) [4 points] We are given $E[x] = 2\theta$. The expectation value of $\hat{\theta}$ is therefore

$$E[\hat{\theta}] = \frac{1}{2n} \sum_{i=1}^{n} E[x_i] = \frac{1}{2n} \sum_{i=1}^{n} 2\theta = \theta \ ,$$

and so the bias is $b = E[\hat{\theta}] - \theta = 0$.

We are given $V[x] = 2\theta^2$. When we take a constant outside of the variance operator, it becomes squared, i.e., $V[\alpha x] = \alpha^2 V[x]$. We are also told that the x_i are independent, so the variance of their sum is the sum of the variances. Using this we find

$$V[\hat{\theta}] = V\left[\frac{1}{2n} \sum_{i=1}^{n} x_i \right] = \frac{1}{4n^2} \sum_{i=1}^{n} V[x_i] = \frac{1}{4n^2} \sum_{i=1}^{n} 2\theta^2 = \frac{\theta^2}{2n} \ .$$

2(c) [4 points] The minimum variance bound (MVB) is given by

$$\text{MVB} = -\frac{\left(1 + \frac{\partial b}{\partial\theta}\right)^2}{E\left[\frac{\partial^2 \ln L}{\partial\theta^2}\right]} \ .$$

We already found the bias b is zero. For the denominator we need the second derivative of $\ln L$,

$$\frac{\partial^2 \ln L}{\partial\theta^2} = \frac{2n}{\theta^2} - \frac{2}{\theta^3} \sum_{i=1}^{n} x_i$$

Taking the expectation value and using $E[x] = 2\theta$ we find

$$E\left[\frac{\partial^2 \ln L}{\partial \theta^2}\right] = \frac{2n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} E[x_i] = \frac{2n}{\theta^2} - \frac{2}{\theta^3}\sum_{i=1}^{n} 2\theta = -\frac{2n}{\theta^2} \; .$$

Putting the ingredients into the equation for the MVB gives

$$\mathrm{MVB} = -\frac{(1+0)^2}{-\frac{2n}{\theta^2}} = \frac{\theta^2}{2n} \; .$$

The MVB is thus the same as the exact variance from (d) and so $\hat{\theta}$ is said to be efficient.

2(d) [4 points] A sketch of the log-likelihood function is shown in Fig. 1; the position of its maximum gives $\hat{\theta}$. The standard deviation $\sigma_{\hat{\theta}}$ is found by moving the parameter away from $\hat{\theta}$ until $\ln L$ decreases by 1/2 from its maximum, as indicated.

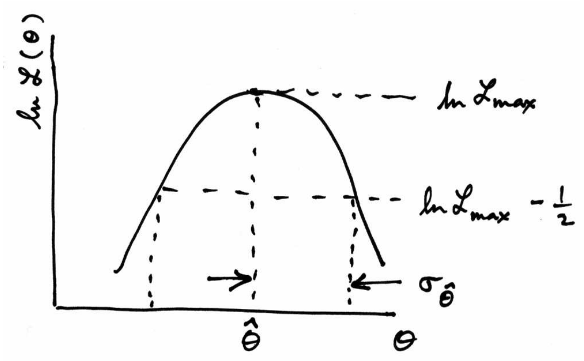

Figure 1: Illustration of finding the ML estimator $\hat{\theta}$ and its standard deviation $\sigma_{\hat{\theta}}$ (see text).

2(e) [4 points] If we regard the number n as a Poisson distributed random variable with mean c/θ, then the full (extended) likelihood function is

$$L(\theta) = \frac{(c/\theta)^n}{n!} e^{-c/\theta} \prod_{i=1}^{n} \frac{x_i}{\theta^2} e^{-x_i/\theta} \; ,$$

and so the log-likelihood is

$$\ln L(\theta) = n \ln \frac{c}{\theta} - \frac{c}{\theta} - 2n \ln \theta - \frac{1}{\theta} \sum_{i=1}^{n} x_i + \text{const.}$$

where the constant represents terms not depending on θ. Setting the derivative of $\ln L$ equal to zero,

$$\frac{\partial \ln L}{\partial \theta} = \frac{n}{\theta} - \frac{c}{\theta^2} - \frac{2n}{\theta} + \frac{n\overline{x}}{\theta^2} = 0 \; ,$$

and solving for θ gives the extended ML estimator

$$\hat{\theta}_{\mathrm{EML}} = \frac{c + n\overline{x}}{3n} \; .$$