Statistical Data Analysis
Problem sheet 9
Due Monday, 14 December 2020

Problem 1: Consider a Gaussian distributed random variable x with mean of zero and a variance v, i.e., the pdf of x is

$$f(x|v) = \frac{1}{\sqrt{2\pi v}} e^{-x^2/2v} \ .$$

Suppose we have a sample of N independent values $x_1, \ldots, x_N$ that follow this pdf, and we wish to make inference about the unknown variance v.

1(a) [**3 marks**] Write down the log-likelihood function $\ln L(v)$ and show that the maximum-likelihood estimator for v is

$$\hat{v} = \frac{1}{N} \sum_{i=1}^{N} x_i^2 \ .$$

1(b) [**2 marks**] Show that the estimator $\hat{v}$ is unbiased.

1(c) [**3 marks**] One can show that the statistic $t = N\hat{v}/v$ follows a chi-squared distribution for N degrees of freedom. Recall also that the chi-squared distribution has a variance of $2N$. Using this information, find the variance of $\hat{v}$.

1(d) [**3 marks**] Using again the fact that $t = N\hat{v}/v$ follows a chi-squared distribution with N degrees of freedom, show that the p-value p_v for a hypothesized value of v can be written

$$p_v = F_{\chi^2_N}\left(N\hat{v}/v\right) \ ,$$

where $F_{\chi^2_N}$ is the cumulative chi-squared distribution for N degrees of freedom. For the p-value, take the set of outcomes with $\hat{v}$ less than the one observed as representing increasing incompatibility with the hypothesized v.

1(e) [**3 marks**] Find the upper limit at confidence level $1 - \alpha$ for v. Express the answer in terms of N, $\hat{v}$ and an appropriate quantile of the chi-squared distribution.

1(f) [**3 marks**] Suppose now we want to use the Bayesian approach to estimate v. Show that the Jeffreys prior $\pi_{\rm J}(v)$ for v is

$$\pi_{\rm J}(v) \propto \frac{1}{v} \ .$$

1(g) [**3 marks**] Using the Jeffreys prior from (f), write down the posterior probability $p(v|x_1, \ldots, x_N)$ as a proportionality. Show that the mode of the posterior is

$$\text{mode}[v] = \frac{1}{N+2} \sum_{i=1}^{N} x_i^2 \ .$$