

Statistics in Computer Science

(Final) Homework Assignment 5

Stanislav Katina, Mojmir Vinkler

December 5, 2016

Exercise 1. Log-normally distributed random variable $X \sim \ln N(\mu, \sigma^2)$ has probability density function

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(\ln(x) - \mu)^2}{2\sigma^2}\right].$$

Assume that σ is fixed and find its Fisher information $\mathcal{I}(\hat{\mu})$.

We have the following random sample distributed log-normally with $\sigma^2 = 1$

4.856, 0.487, 0.580, 0.839, 0.721, 2.416, 0.715, 0.361, 0.703, 5.829

Plot scaled log-likelihood function of the sample with μ on x -axis and scaled log-likelihood $l^*(\mu|\mathbf{x}) = l(\mu|\mathbf{x}) - \max(l(\mu|\mathbf{x}))$ on y -axis. Compare $l^*(\mu|\mathbf{x})$ with its quadratic approximation calculated using Taylor approximation $\ln\left(\frac{L(\mu|\mathbf{x})}{L(\hat{\mu}|\mathbf{x})}\right) \approx -\frac{1}{2}\mathcal{I}(\hat{\mu})(\mu - \hat{\mu})^2$.

Exercise 2. Load dataset one-sample-mean-skull-mf.txt with variable skull.L that represents skull length of ancient Egyptian male population in mm. Let's assume that the skull length is normally distributed, i.e. $N(\mu, \sigma^2)$.

1. Test null hypothesis that mean of skull length is equal to 177.568 mm ($H_0 : \mu = 177.568$) on significance level $\alpha = 0.05$.
2. Calculate empirical $100 \times (1 - \alpha)\%$ confidence interval for population mean of skull length with $1 - \alpha = 0.95$.

For (1) use Wald test statistic T_W and likelihood ratio test statistics U_{LR} . For (2) use empirical Wald confidence interval and empirical likelihood confidence interval.

Exercise 3. Statistical graphics

1. Take some figures combining raster graphics and statistical graphics from the press - (A) three good examples and (B) three bad examples and describe the reasoning behind your choice, i.e. why the figure is good or bad. See PDF Statistical Graphics in Czech and World Press for an inspiration.
2. Take one from good or bad examples from above assignment and use the numbers to create your own figure. This figures should look better than an example from the press.

3. Create figure in \mathbb{R} as superimposition of raster graphics and statistical graphics. Describe what can be seen on the figure. To import any raster graphics to \mathbb{R} and its processing, use one or more of the following libraries - rasterImage, library(png), library(pixmap), library(rimage), library(tiff) library(raster), library(rasterVis), library(maps), library(mapsdata) or library(mapstools).