Stat 314, Problem Set 1
Autumn 2013/14

Due in class on paper, Thursday October 17, 2013
This one is individual effort; later ones will feature team work.

1. Problems from the text: 2.2, 2.3, 3.1, 3.3, 3.4, 3.7

2. The file `ducks.txt` has data on 11 ducks. These ducks were a cross between mallards and pintails. Their plumage and behavior were both rated on scales, with higher values being more pintail-like.

Compute the empirical likelihood for the mean of these data on a grid of points covering the rectangle $[4, 15] \times [3, 15]$. This is a bounding box for data and it includes many points outside of the convex hull.

You may use the empirical likelihood R code in the course web page. For data in the matrix $Z \in \mathbb{R}^{n\times d}$, calling `emplik(Z,mu)` will test whether they are IID with mean mu $\in \mathbb{R}^d$. The default mu is 0. So you can also call `emplik(ZZ)` where ZZ has had mu subtracted from each row.

Values outside of the convex hull will give super small log likelihoods, such as -700 or less.

Plot contours of empirical likelihood corresponding to 90, 95, 99 and 99.9 percent confidence levels. Superimpose the actual data. Hint: it is tricky to get contour and image plots just right. Using a $k \times (k+1)$ grid will expose errors you might make that a square $k \times k$ grid could hide.

3. Bootstrap the duck data 1000 times. On each bootstrap sample record the empirical log likelihood of $\bar{X}$ the mean in the original sample. Make a QQ plot comparing the bootstrap values $-2\log\mathcal{R}^*(\bar{X})$ to the $\chi^2_{(2)}$ distribution.

4. For the worm data plot contours of the empirical likelihood for pairs (μ, σ) where μ is the hypothesized mean and σ^2 is the hypothesized variance.

Show contours for 50, 90, 95 and 99 percent confidence using a χ^2 calibration. It takes some exploration to find the relevant range.

For normally distributed data the estimated mean and variance are statistically independent. But these data are skewed.