HW 2 (IQB w18)

19/1/18

1. Information of altered gene expression levels (such as for certain proteins in the apoptotic death pathway) in cancer cells provides crucial insights into the generation of malignant phenotype. Carry out simple statistical data analysis of gene expression data
   http://cancer.sanger.ac.uk/cosmic/cnv/details?ln=BCL2&expn=over
   (obtained from the COSMIC database)

(a) (1 pt) Define gene expression level and z score. Mention an experimental technique for assessing gene expression levels.

(b) (1 pt) Generate probability distribution (histogram plot) of expression level (z scores) for the anti-apoptotic protein Bcl2. Estimate average and standard deviation in z values (use data for at least 20 samples).

(c) (1 pt) Estimate statistical correlation between copy number variation (cnv) and gene expression levels for Bcl2. Explain your results.

   For 1(c) use the data provided in the column labeled as Copy Number.