Exercise 1 (Poisson distribution)

During measurements through Positron emission tomography (PET) a detector produces "clicks". These clicks are counted for each volumetric pixel (voxel). The counts of different voxels (v1,v2) from different PET machines (a,b,c) are stored in the dataset. The data can be downloaded from the website (**R-Hint**: `read.table(..., sep=",", header=TRUE)`).

(a) Since the data contains count data (number of "clicks" per voxel), we fit a poisson model. The poisson model is specified by the value of the parameter λ (Lambda).

- Estimate the parameter λ by calculating the mean over the counts of `a.v1`.
- Plot the distribution of the counts, that you would expect from the poisson model (**R-Hint**: `plot(dpois(0:50,lambda=...),...)`)

(b) Find out how well the poisson model describes the data. Plot the distribution of the observed data `a.v1` as histogram together with the poisson distribution. Additionally use a QQ-Plot to compare the expected to the observes quantiles of the counts. (**R-Hint**: `hist(..., probability=TRUE)`, `lines(dpois(...))`. For the qqplot use the function `qqPlot()` from library `car`. If the library is not installed used `install.packages("car")` and `library(car)`. If the library is installed, you can load it directly with `library(car)`)

(c) The literature on PET states that a simple poisson model is not always well suited due to overdispersion, i.e. the variance in the data is higher than assumed in the poisson model. Consider the mean and the variance for the other three variables (`a.v2, b.v1, c.v1`). Do you think simple poisson models are appropriate? Verify your assumptions with QQ-plots.

Exercise 2 (Binomial distribution)

Imagine that there is a new treatment offering a relief for 70% of the patients. This treatment is given to 100 new patients.

(a) What is the probability that the symptoms disappear in exactly 60 patients (**R-Hint**: `dbinom()`)?

(b) Simulate 300 data points for the situation described above (**R-Hint**: `rbinom()`)?

(c) Visualize both the simulated data and the underlying theoretical distribution.

Email: hezo@zhaw.ch