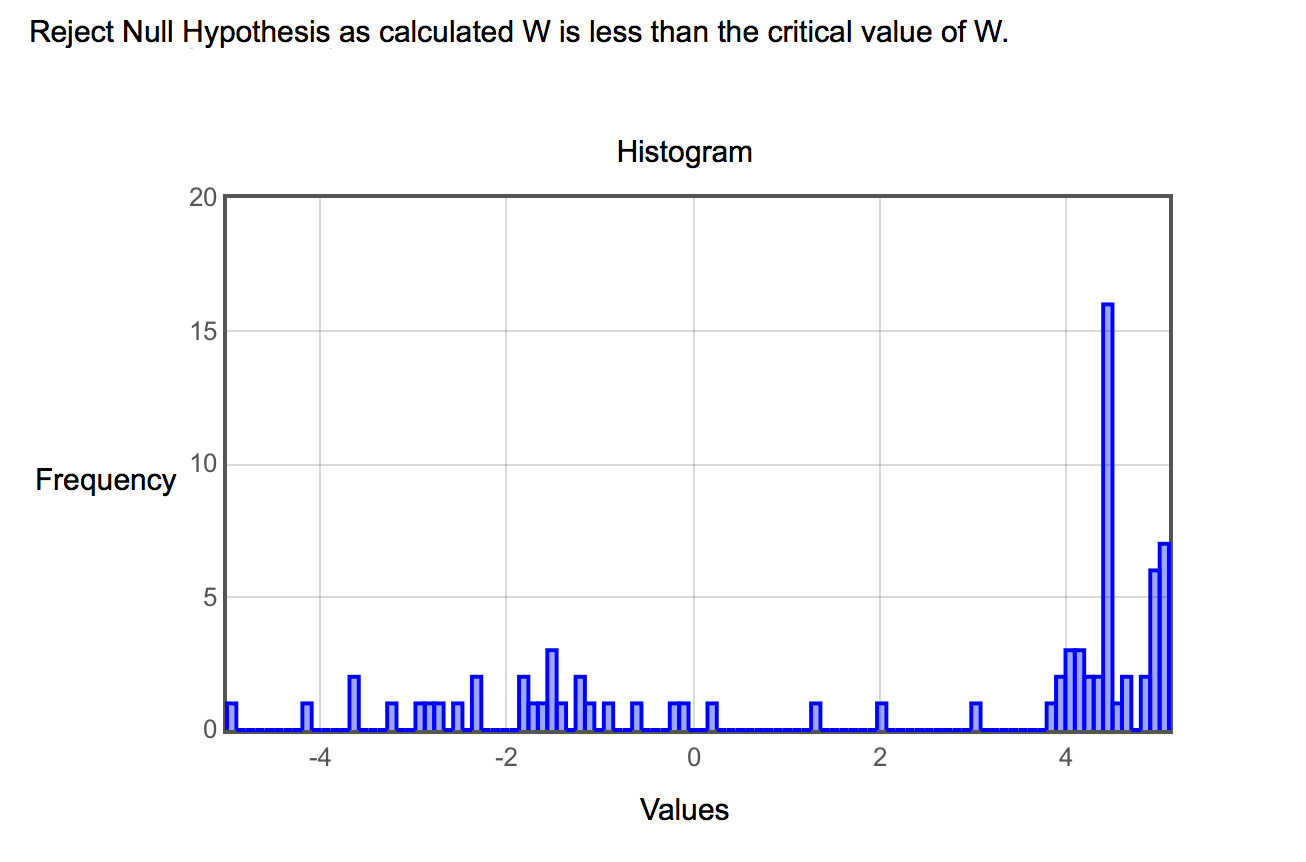
1. **Analysis of CMTR1 and CMTR2 against YFV results using the RNA as reference molecule**

Here I perform 4 different dockings but using the capped RNA found in the PDB as reference molecule. The goal is to observe if the best ligands found in YFV are stable in the RNA binding pocket.

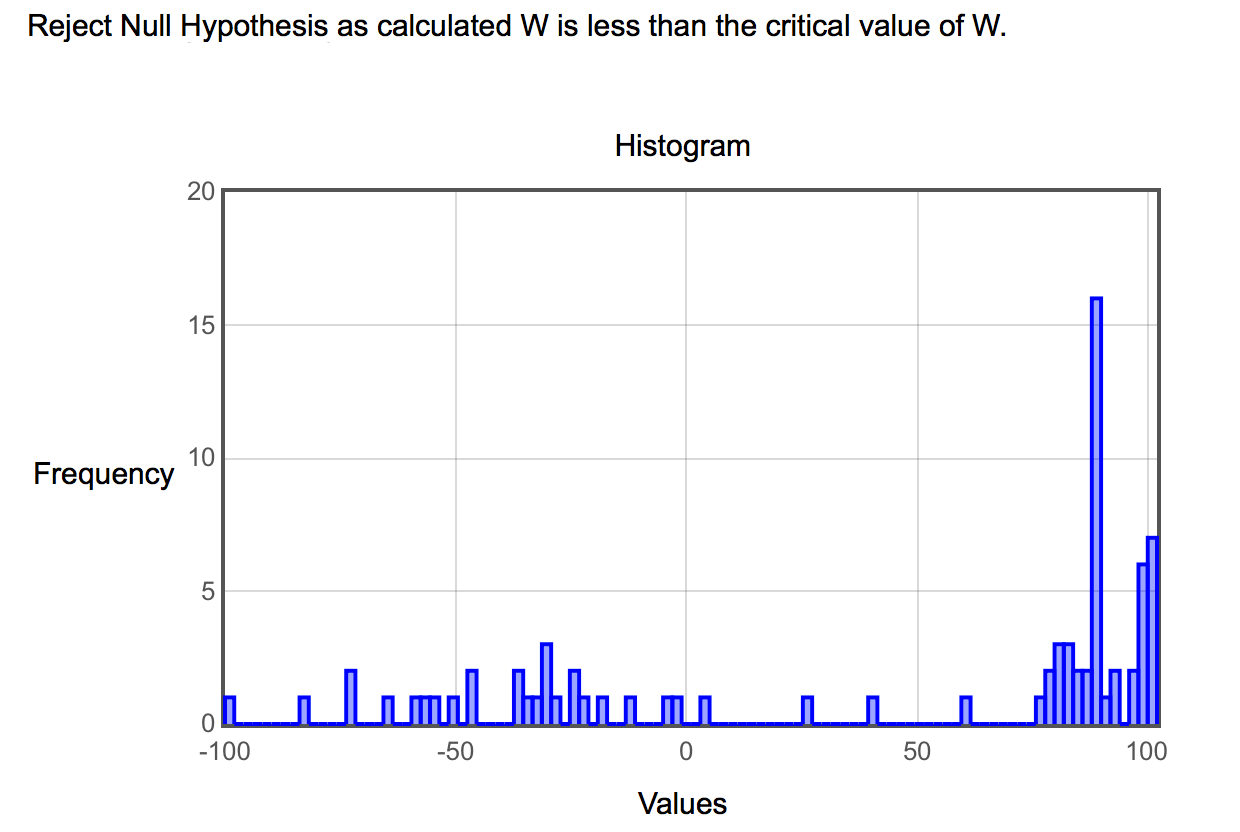
* **Kolmogorov-Smirnov Comparison of Two Data Sets RESULTS**

| Receptor/Binding site | 3EVA\_sah ligands (75) | 3EVC\_gtp ligands (77) |
| --- | --- | --- |
| CMTr1/SAM | p-value > 0.05, no differences found | **p-value < 0.05, differences found** |
| CMTr2/SAM | **p-value < 0.05, differences found** | **p-value < 0.05, differences found** |
| CMTr1/RNA | p-value > 0.05, no differences found | p-value > 0.05, no differences found |
| CMTr2/RNA | p-value > 0.05, no differences found | p-value > 0.05, no differences found |

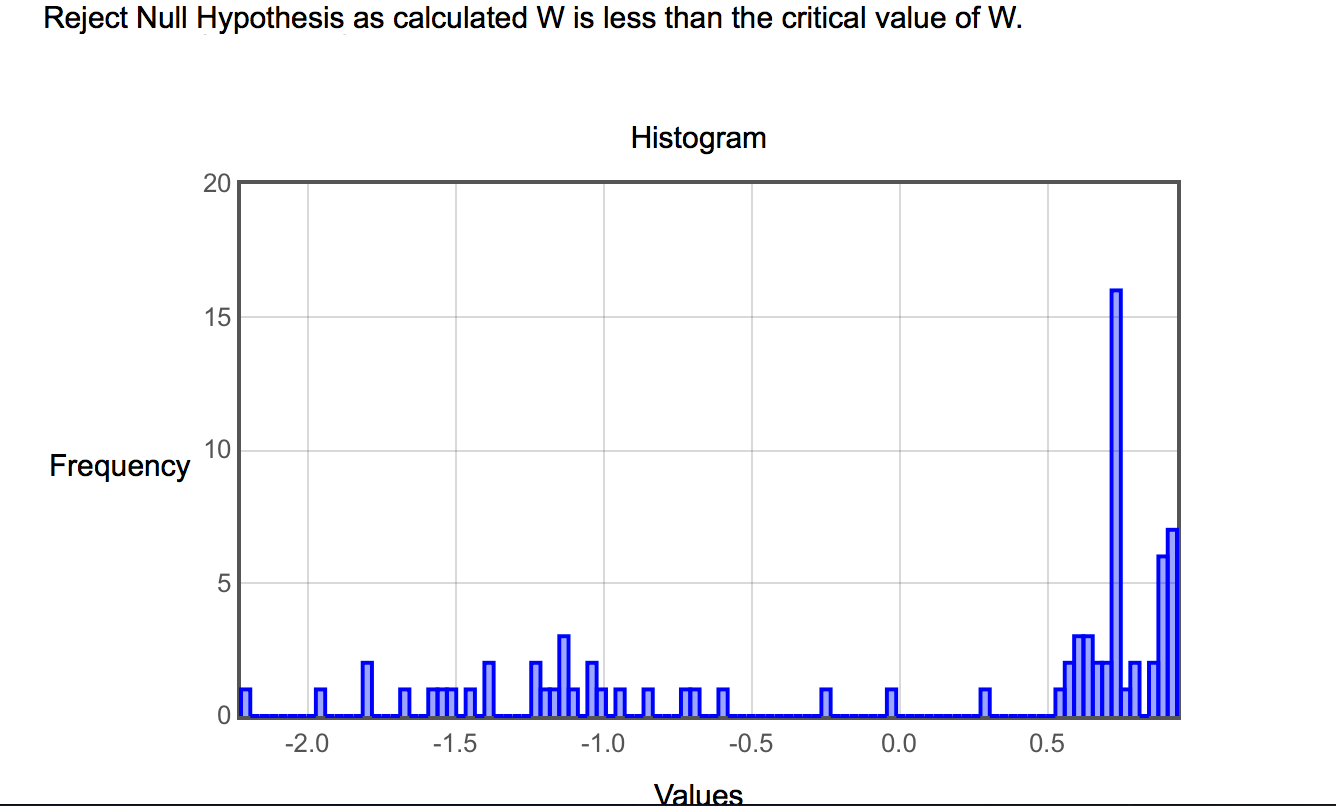
* **Using Normalizer in KNIME (Min-Max Normalization [-5, 5])**



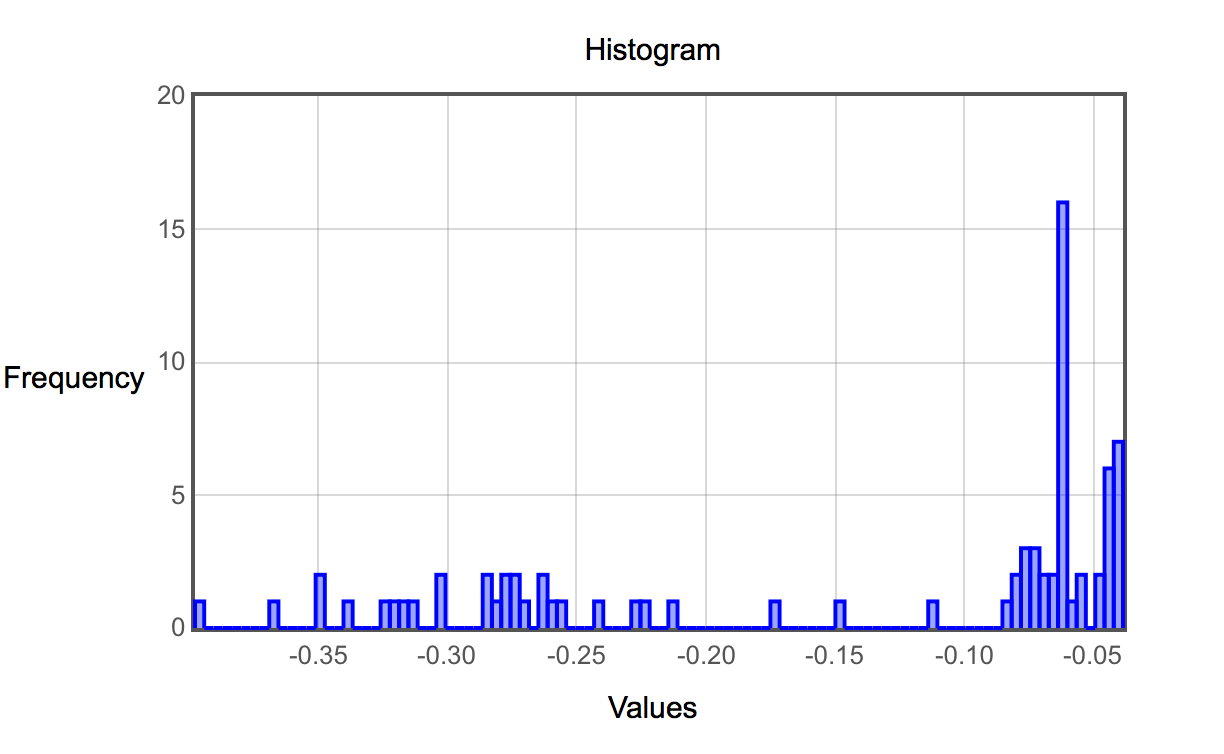
* **Using Normalizer in KNIME (Min-Max Normalization [-100, 100])**



* **Using Normalizer in KNIME (Z-Score Normalization, Gaussian)**

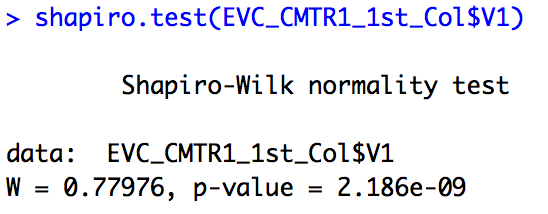


* **Using Normalizer in KNIME (Normalization by Decimal Scaling)**

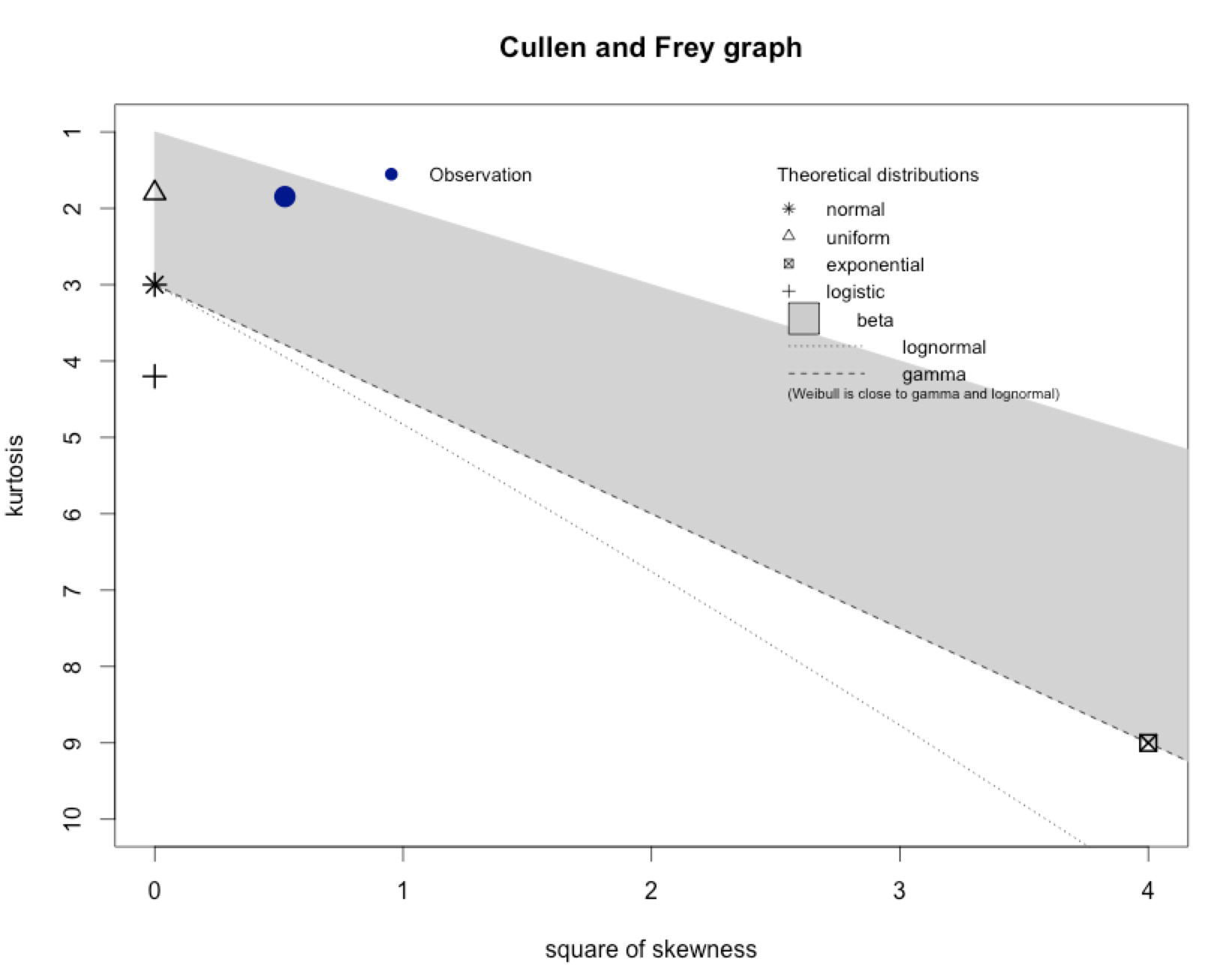


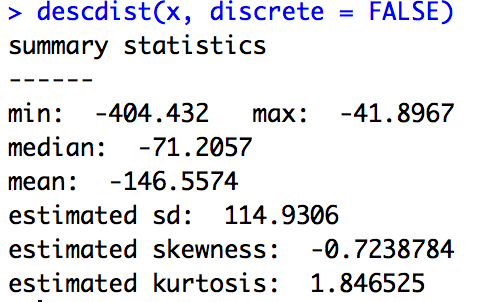
All Shaphiro-Wilk tests reveals that the distribution of these datas are not following a normal distribution (in fact is skewed to the right), so most of the results are concentrated in the right side (more negative values). On this case we can’t apply a parametric test like the t-student statistical test -> **Reject Null Hypothesis as calculated W is less than the critical value of W, so distribution is not following a normal distribution.**

**R: EVC\_CMTR1\_1st\_Col <- read.csv2("3EVC\_CMTR1\_1\_n.csv", header = FALSE)**

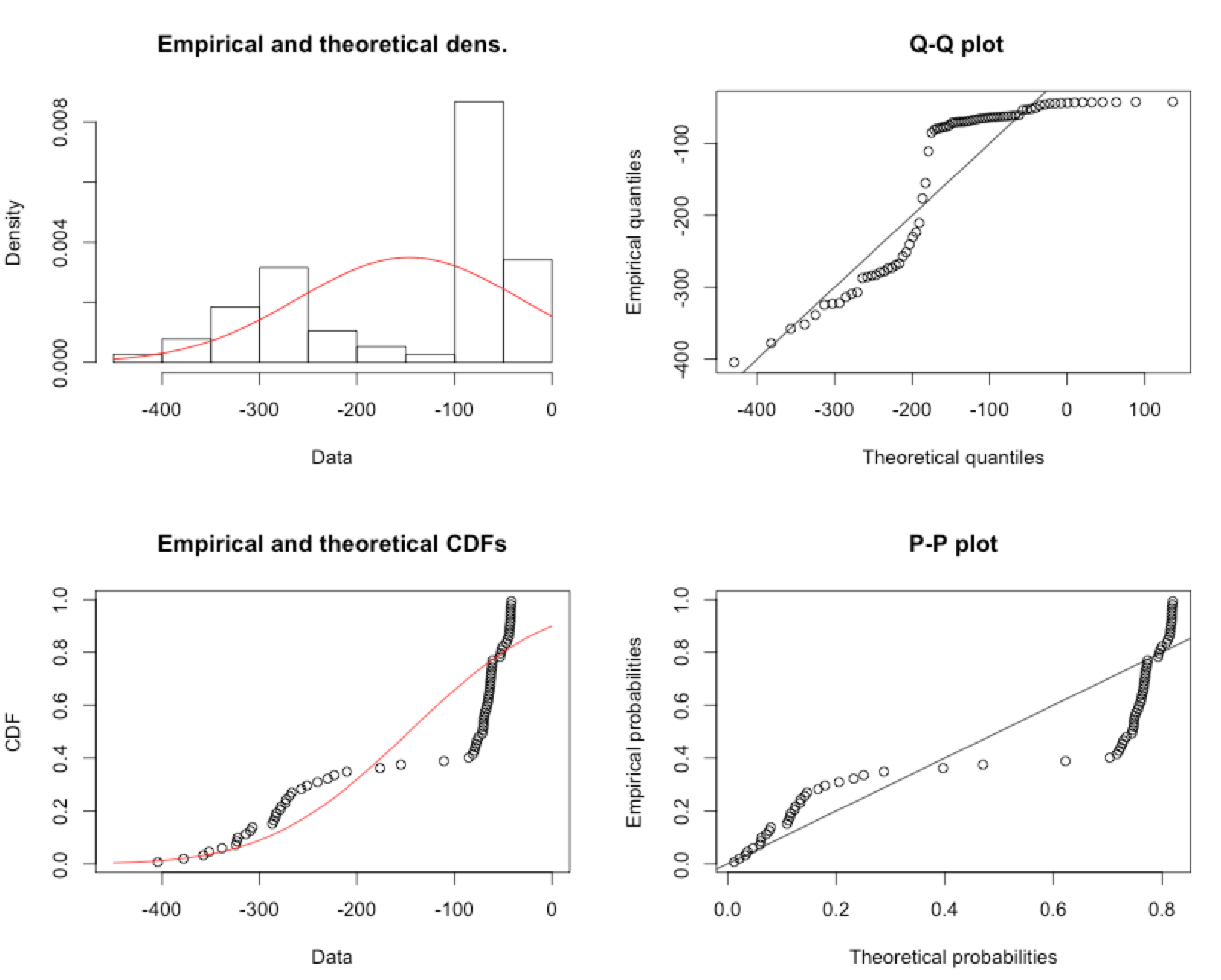


1. **Analysis of CMTR1 and CMTR2 against YFV results using the RNA as reference molecule (USING R and non-normalized data)**

Here I’m analyzing the data with some R commands and using library(fitdistrplus). We want to figure out which distribution is our data following, and as usual, if it’s skewed or not.

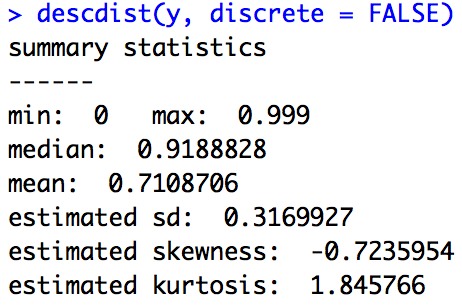
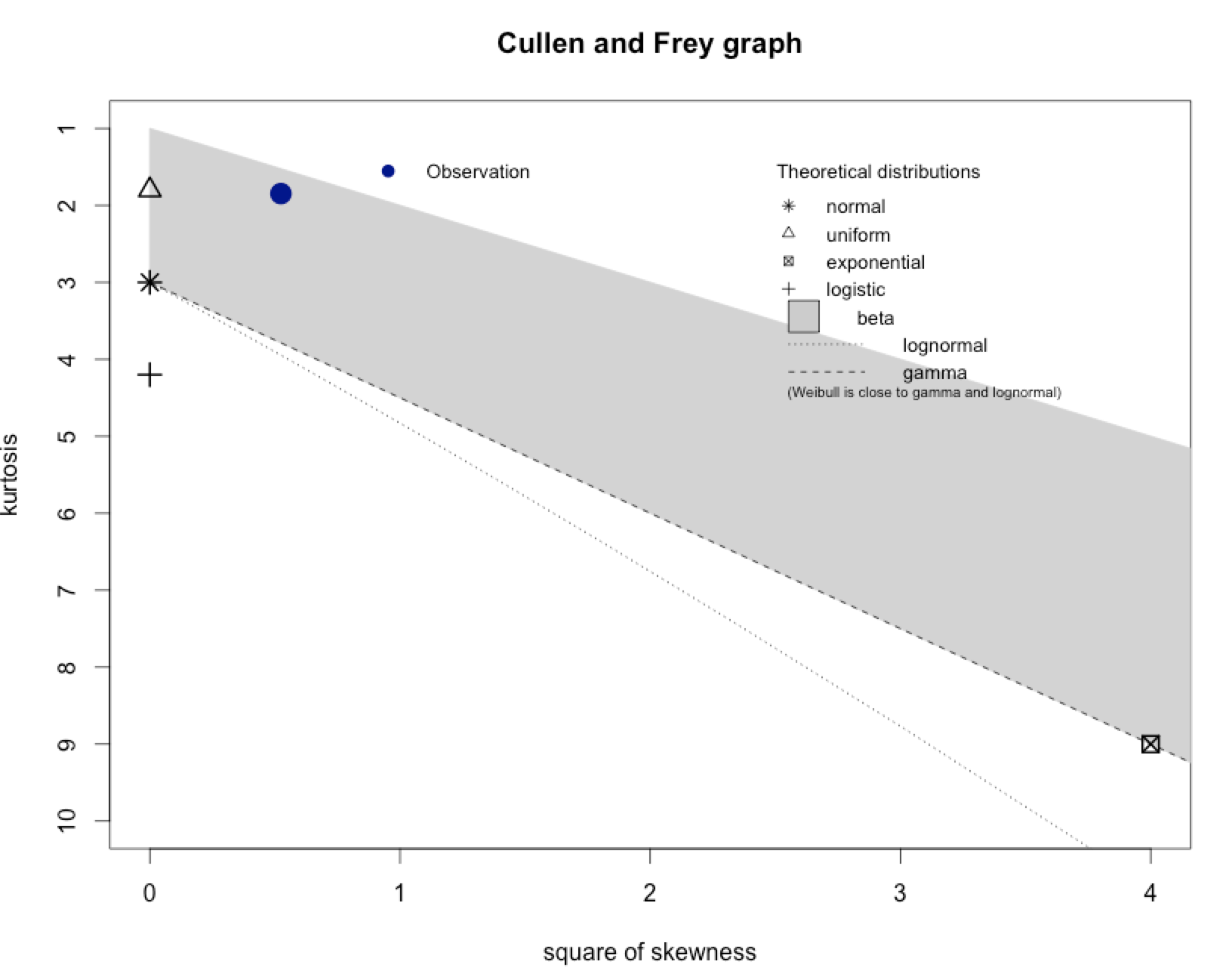


The kurtosis and squared skewness of my sample is plotted as a blue point named "Observation". It seems that the possible distributions included the beta distribution.



1. **Analysis of CMTR1 and CMTR2 against YFV results using the RNA as reference molecule (USING R and normalized data)**

The kurtosis and squared skewness of my sample is plotted as a blue point named "Observation". It seems that the possible distributions included the beta distribution.



I had to remove the 0.0 value because it gave me problems: Error in fitdist(y, distr = "beta") : the function mle failed to estimate the parameters, with the error code 100.

> fit.beta <- fitdist(y, distr=“beta")

