**Homework #2: Maximum likelihood estimation (Two\_Parameter Case)**

1. In HW #1, we consider which a sample of , each of which follow a Weibull distribution defined as



for γ and β >0.

1. Please derive the score equations of (γ, β).
2. Then choose a test statistics that you learn in the class to test γ = 2. (There are three options) Please state the decision rule based on the test.

Please apply the formula that you derive above to fit the wind speed data:

9.4 12.7 3.9 9.8 9.5 15.0 8.1 15.7 7.8 9.1 14.5 10.3 9.0 11.1 5.2

3.0 5.7 10.1 7.6 17.7 6.4 9.1 12.9 9.6 6.2 8.6 17.0 17.0 16.8 5.1

8.6 20.1 4.6 7.3 14.3 8.7 8.8 4.4 8.2 10.8 5.3 5.5 10.8 21.9 4.4

6.9 5.2 6.9 12.2 11.6 16.0 16.8 9.9 7.3 4.6 3.2 2.4 5.1 5.6 3.5

6.4 3.9 9.9 7.7 10.3 15.9 17.3 12.7 10.0 9.0 16.2 12.2 6.6 11.2 14.3

14.5 13.0 11.9 8.8 10.0 18.5 23.2 21.5 8.9 6.4 9.9 7.9 8.0 12.9 8.3

6.5 9.8 12.9 16.2 5.9 11.4 4.6 7.2 12.5 5.3 15.8 5.5 10.9 8.6 15.0

10.1 7.8 11.8 10.7 12.6 11.5 16.1 5.9 6.3 9.6 19.2 13.9 6.9 7.6 5.1

10.3 4.1 15.2 9.8 9.9 11.4 6.9 15.6 5.8 10.7 15.0 6.8 6.0 8.4 10.4

* 1. 7.8 21.6 14.9 5.0 18.5 10.3

1. What are the fitted MLE values of (γ, β)?
2. Do the data support γ = 2?
3. Please draw a Q-Q plot (X-axis: the data quantiles; Y-axis: the quantiles of the Raleigh distribution).
4. Please comment on the results of (d) and (e).

**2:** It is conjectured that *tropical plants of a certain genus tend to produce more flowers at higher altitudes than at lower altitudes.* Fifteen species in this genus are known to occur at both altitudes in a particular country.

To test the conjecture, one plant from a lowland forest and one plant from higher altitudes were collected from each of twelve species from this genus.

The number of flowers on each plant were counted, and the results were:

Species LowAlt HighAlt Species LowAlt HighAlt

1 5 19 7 3 17

2 4 10 8 4 14

3 12 4 9 6 3

4 7 10 10 15 3

5 17 17 11 9 9

6 4 12 12 7 10

1. Perform a sign test based on the paired data (LowAlt,HighAlt).

Compute the p-value. Note: the hypothesis is one-sided.

1. Perform a (one-sided) signed-rank test based on the normal approximation. Compute the p-value.
2. Perform a signed rank test using a software (based on the exact distribution)
3. Even though the data is from 24 different plants, why would it be incorrect to assume that the plants from the lowlands and the plants from higher altitude form two independent samples? (為何不用雙樣本? 想一想. 沒有標準答案)

3

The data in the file “Soil.csv” represent soil water content (% water by volume) for independent random samples of soil from two experimental fields growing bell peppers.

The water content of soil (%) collected from two places are under comparison. Place 1 contains 72 observations with the average water content around 11.42%; Place 2 contains 80 observations with the average water content around 10.65%.

1. Please plot the two empirical distribution functions and compare them.
2. Please perform a two-sample t test for testing  under 5% level of significance and state your conclusion. Comment on whether it is appropriate to use the t test.
3. Please perform a Wilcoxon rank sum test for testing  under 5% level of significance. State your conclusion.