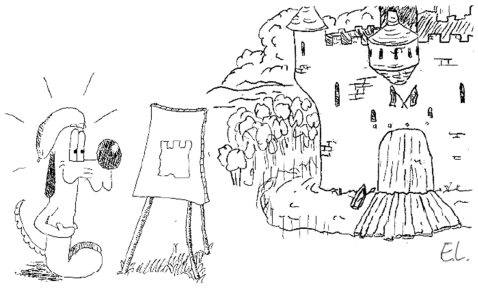
PERFORMANCE EVALUATION EXERCISES

MODEL FITTING 2

Jean-Yves Le Boudec, Spring 2021



1. We have a data setx_i, $i = 1 : n$ with n large, for which we found that the hazard rate $\lambda(x)$ becomes small when x is large. From this observation, which of the following distributions could be envisioned to model the data ?

   (a) ☐ A normal distribution

   (b) ☐ A Weibull distribution with shape parameter $0 < c < 1$

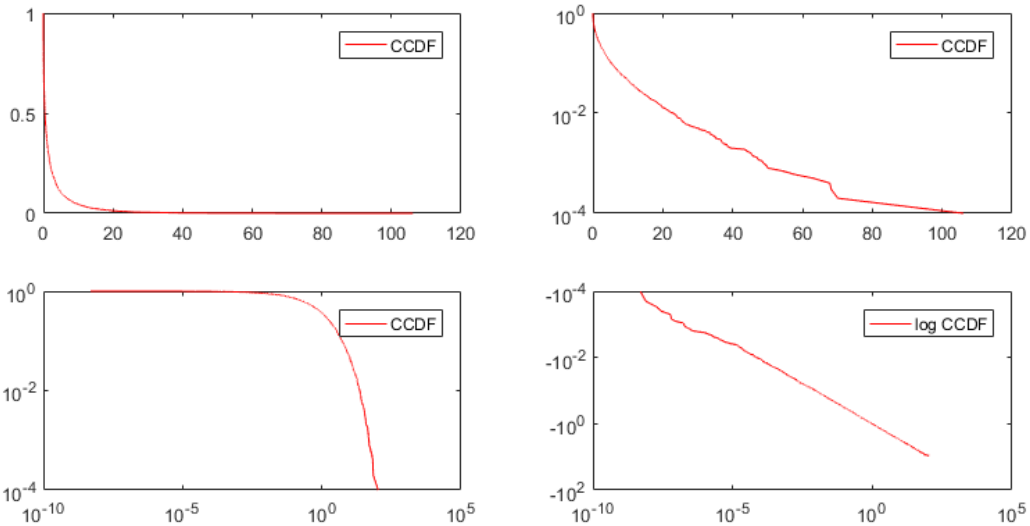
   (c) ☐ An exponential distribution

   (d) ☐ A Weibull distribution with shape parameter $c > 1$

   (e) ☐ A Pareto distribution with index $0 < p < 2$

   (f) ☐ A Pareto distribution with index $p \geq 2$

2. (Continued) We plot the survival function of the data set in the previous question (i.e. the complementary CDF (CCDF)). We obtain the following plots in various scales.



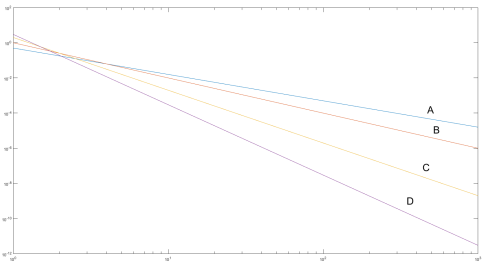
   Which distribution do you propose to model this data set ?

   (a) ☐ A Pareto distribution with $0 < p < 2$

   (b) ☐ A Pareto distribution with $2 \leq p$

   (c) ☐ A Weibull distribution

3. Find the index of each of the standard Pareto PDFs shown in the figure



   (a) ☐ $A = 0.5, B = 1, C = 2, D = 3$

   (b) ☐ $A = 3, B = 2, C = 1, D = 0.5$

   (c) ☐ $A = 0.5, B = 2, C = 1, D = 3$

   (d) ☐ $A = 1, B = 2, C = 3, D = 0.5$

4. The complementary CDF of a Pareto distribution follows a power law...

   (a) ☐ True

   (b) ☐ False

   (c) ☐ It depends on the index p

5. A Pareto distribution is heavy tailed (i.e. with infinite variance)...

   (a) ☐ True

   (b) ☐ False

   (c) ☐ It depends on the index p

6. For a Pareto distribution, the hazard rate $\lambda(t)$ is such that $\lim_{t\to\infty} \lambda(t) = 0$.

   (a) ☐ True

   (b) ☐ False

   (c) ☐ It depends on the index p

7. The distribution of the sum of n iid random variables with heavy tail and index $p < 2$, for large n, is approximately...

   (a) ☐ Normal

   (b) ☐ Stable with same index p

   (c) ☐ Stable but not necessarily with same index p

   (d) ☐ Poisson

   (e) ☐ It depends on p

8. X is a random variable with distribution standard Pareto with index $p > 0$. The distribution of $\log(X)$ is ...

   (a) ☐ Normal

   (b) ☐ Stable with same index p

   (c) ☐ Stable but not necessarily with same index p

   (d) ☐ Poisson with rate $\lambda = \frac{1}{p}$

(e) ☐ Poisson with rate $\lambda = p$

(f) ☐ Exponential with rate $\lambda = p$

(g) ☐ Exponential with rate $\lambda = \frac{1}{p}$

(h) ☐ Lognormal

9. The distribution of the sum of n iid random variables with finite variance, for large n, is approximately...

   (a) ☐ Normal

   (b) ☐ Stable

   (c) ☐ Poisson

   (d) ☐ It depends

10. X_i is an iid sequence with PDF $f_{X_i}(x) = \frac{2}{\pi(1+x^2)}\mathbf{1}_{\{x\geq 0\}}$ (one-sided Cauchy). When n is large, what can you say about $Y = X_1 + ... + X_n$?

    (a) ☐ It is approximately gaussian

    (b) ☐ It is approximately stable with index $p = 2$

    (c) ☐ It is approximately stable with index $p = 1$

11. How do you generate a sample of the standard Weibull distribution with shape parameter c ?

12. What are the models and the null hypothesis of a Jarque-Bera test ? Give a formula to compute the $p-$value when the data is $x_1, ..., x_n$ and n is large.

13. (a) What is the complementary CDF of a Pareto distribution with index p rescaled by a factor $s > 0$?

    (b) What is the distribution of a censored standard Pareto random variable, more precisely, the conditional distribution of a standard Pareto random variable X given that $X > a$?

    (c) A data set $X_1, ...X_n$ is assumed to be iid from a censored standard Pareto with index p and censoring parameter a. Write the formula for the maximum likelihood estimation of p and a.