Goodness of fit

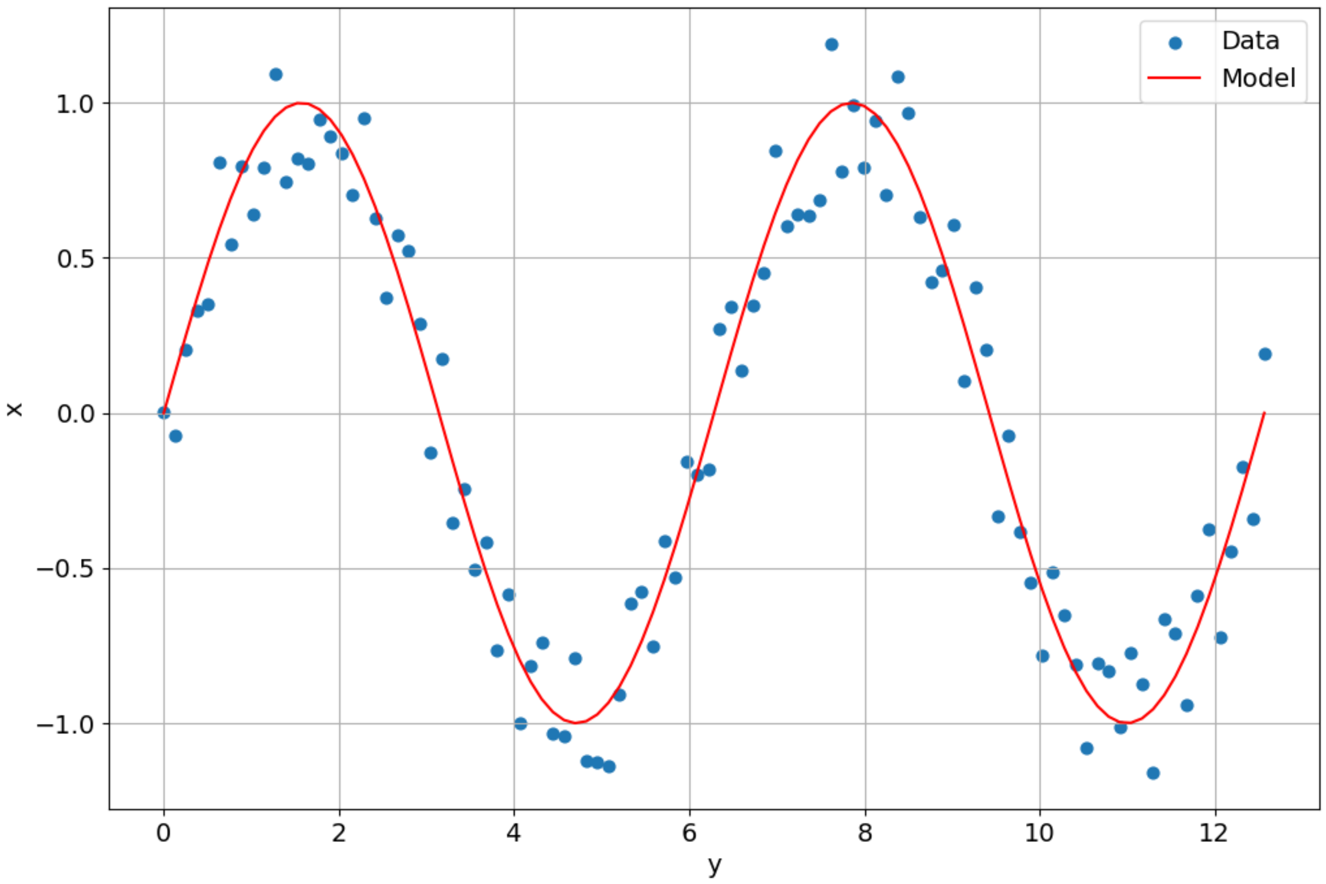
Oswaldo Morales Napoles

Patricia Mares Nasarre

TU Delft

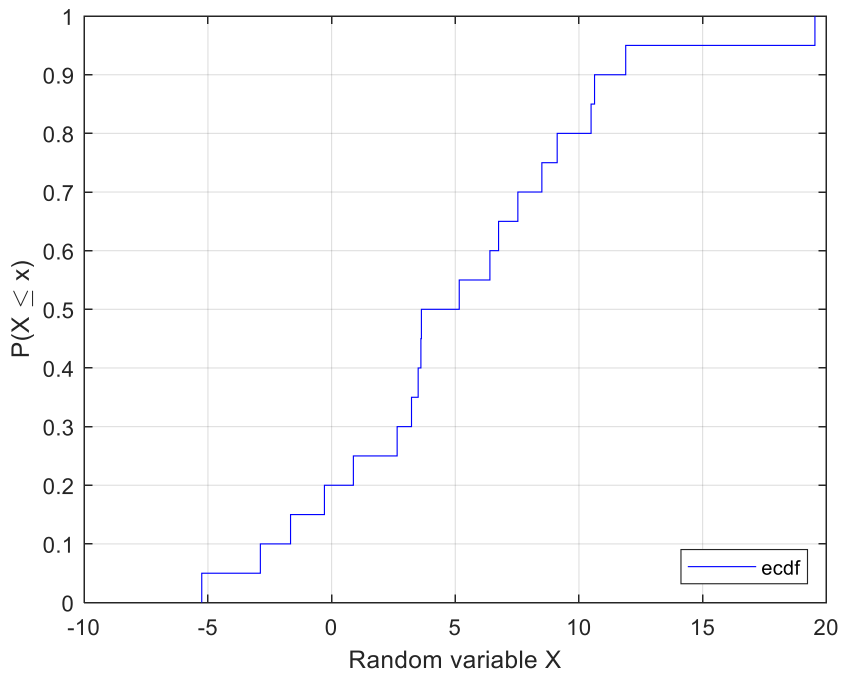
Goodness of fit

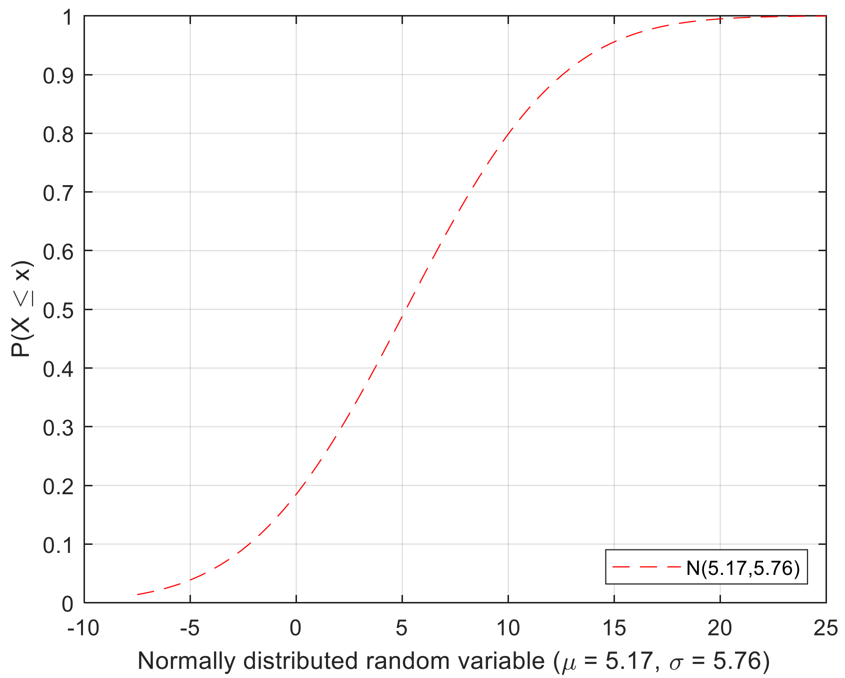
- Fit a statistical model to data
- See for example likelihood lecture
- How “good” is our model?
- Assess through **goodness of fit** techniques
- Kolmogorov-Smirnov (KS) test
- Graphical methods (not in detail in this lecture)



KS statistic

- $D_n = sup_x\,|\hat{F}(x) - F(x)|$
- $\hat{F}(x)$ is the empirical cumulative distribution
  - (# samples ≤ x) / n
- $F(x)$ is some parametric cumulative distribution
- The statistic is (roughly) the maximum distance between $\hat{F}(x)$ and $F(x)$





TU Delft

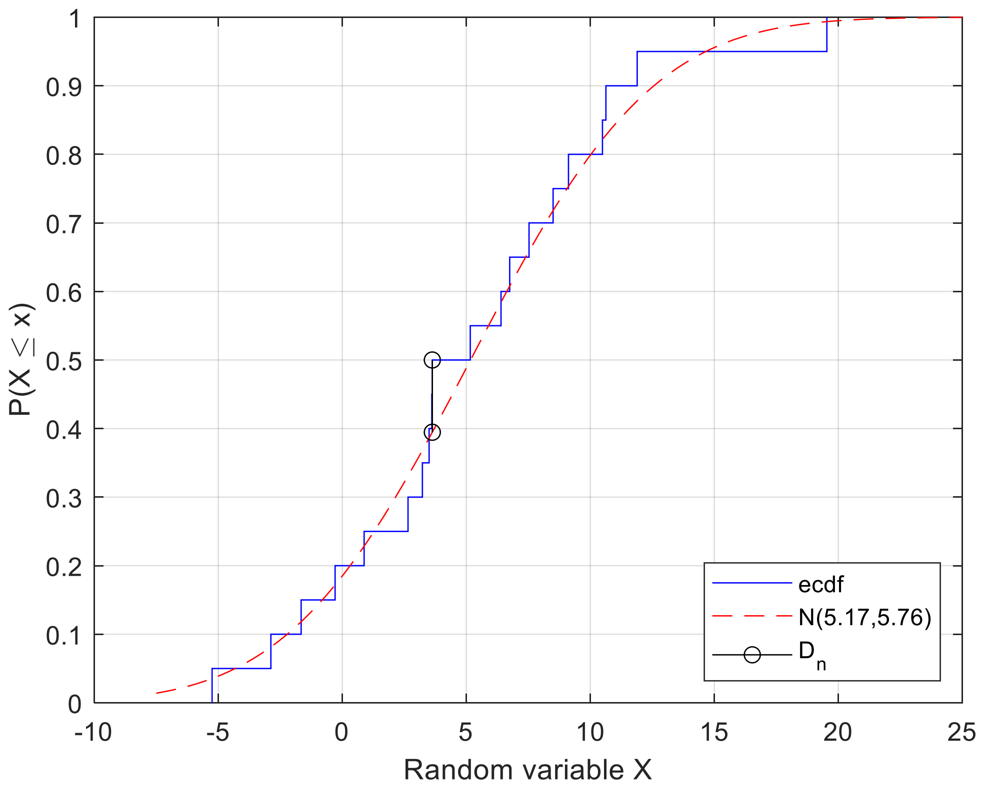
KS statistic (example)

- $X = \{19.54; 9.12; 11.89; -0.29; 2.65; 3.63; 10.49; 3.61; 8.50; -5.25; \ldots$
  $3.23; 0.88; -2.88; 7.53; 6.40; 5.16; -1.66; 10.63; 6.75; 3.50\}$

- $D_n = sup_x\, |\hat{F}(x) - F(x)| \approx 0.1054$

- Test statistic not enough.

- Formal hypothesis test
  - $H_0: \hat{F} \sim F$ ($\hat{F}$ has the same distribution as F)



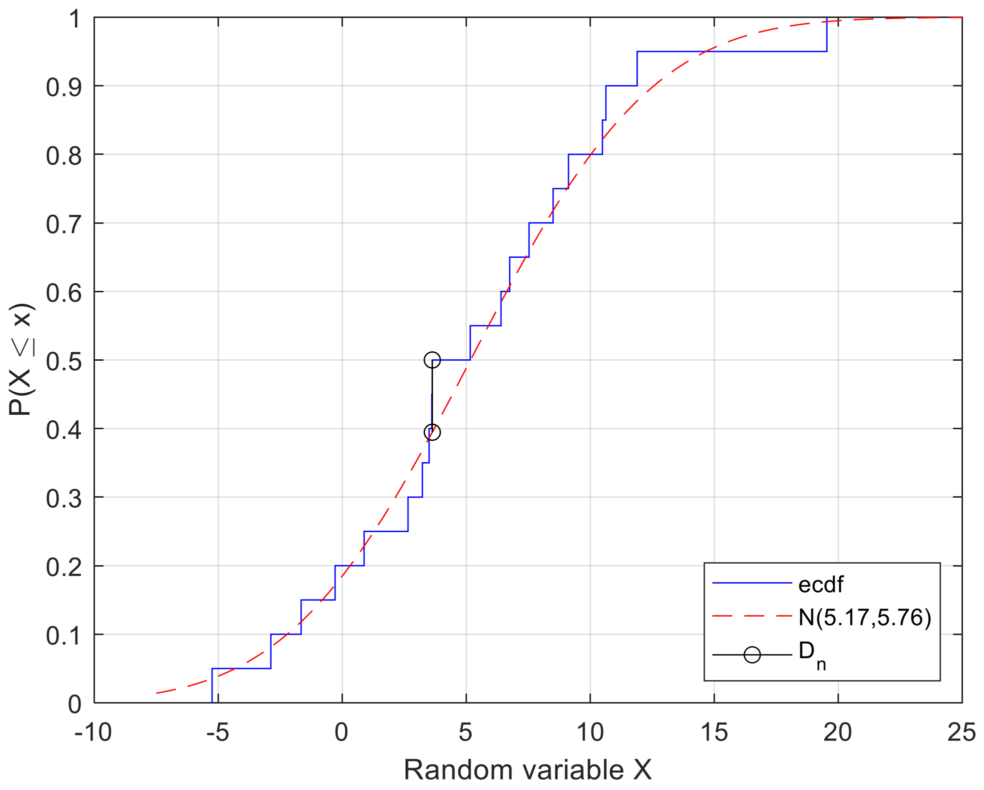
TU Delft

KS statistic (example)

- Formal hypothesis test
  - $H_0: \hat{F} \sim F$ ($\hat{F}$ has the same distribution as F)

- The distribution of D_n is required
  - Implemented in statistical software
  - For different cases

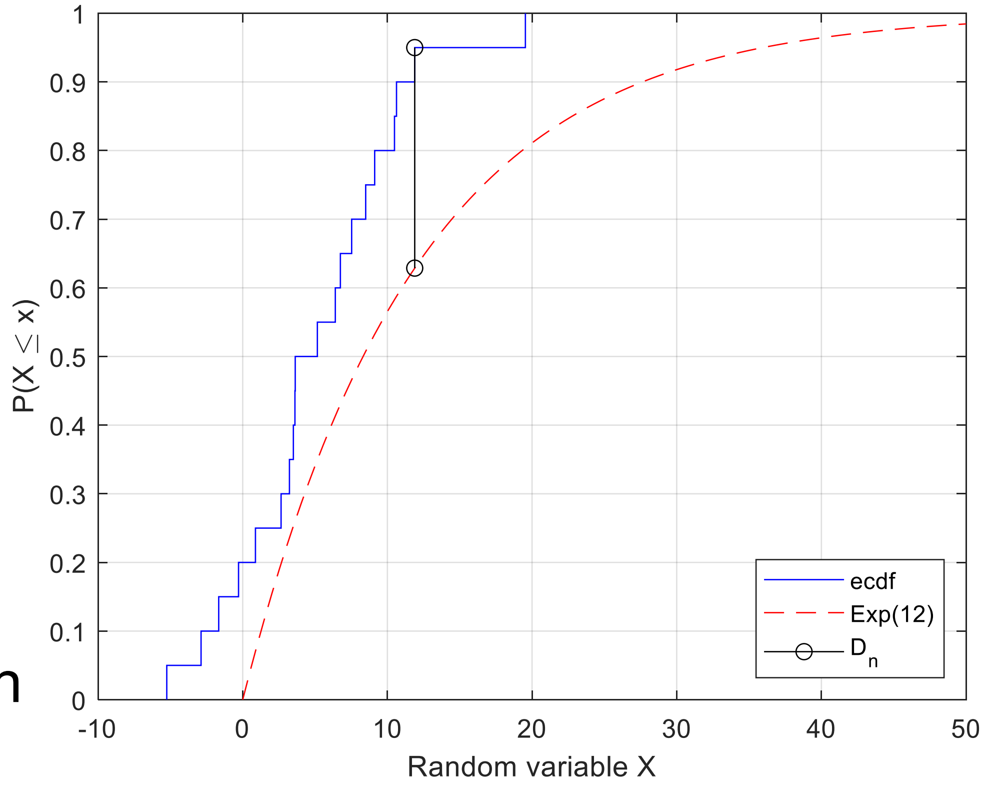
- We cannot reject H_0 for a $N(5.17, 5.76)$



TU Delft

KS statistic (example)

- Formal hypothesis test
  - $H_0 : \hat{F} \sim F$ ($\hat{F}$ has the same distribution as F)
- The distribution of D_n is required
  - Implemented in statistical software
- **Can** reject H_0 for an $Exp(12)$
  - No reason to believe sample comes from an $Exp(12)$ distribution



TUDelft

KS remarks

- Formal hypothesis test
- Relatively easy to understand (Relatively intuitive)
- Widely used & implemented in statistical software
- Many other methods available. For example:
  - Graphical GOF techniques

Graphical procedures

- Graphical procedures:
  - data presentation
  - confirmation of analysis

Source: https://monsterwriterblog.wordpress.com/category/how-to-perform-visual-assessment/

Graphical procedures

- Graphical procedures:
  - data presentation
  - confirmation of analysis

- Two techniques here:
  - QQ-plot
  - Probability plot

Source: https://monsterwriterblog.wordpress.com/category/how-to-perform-visual-assessment/

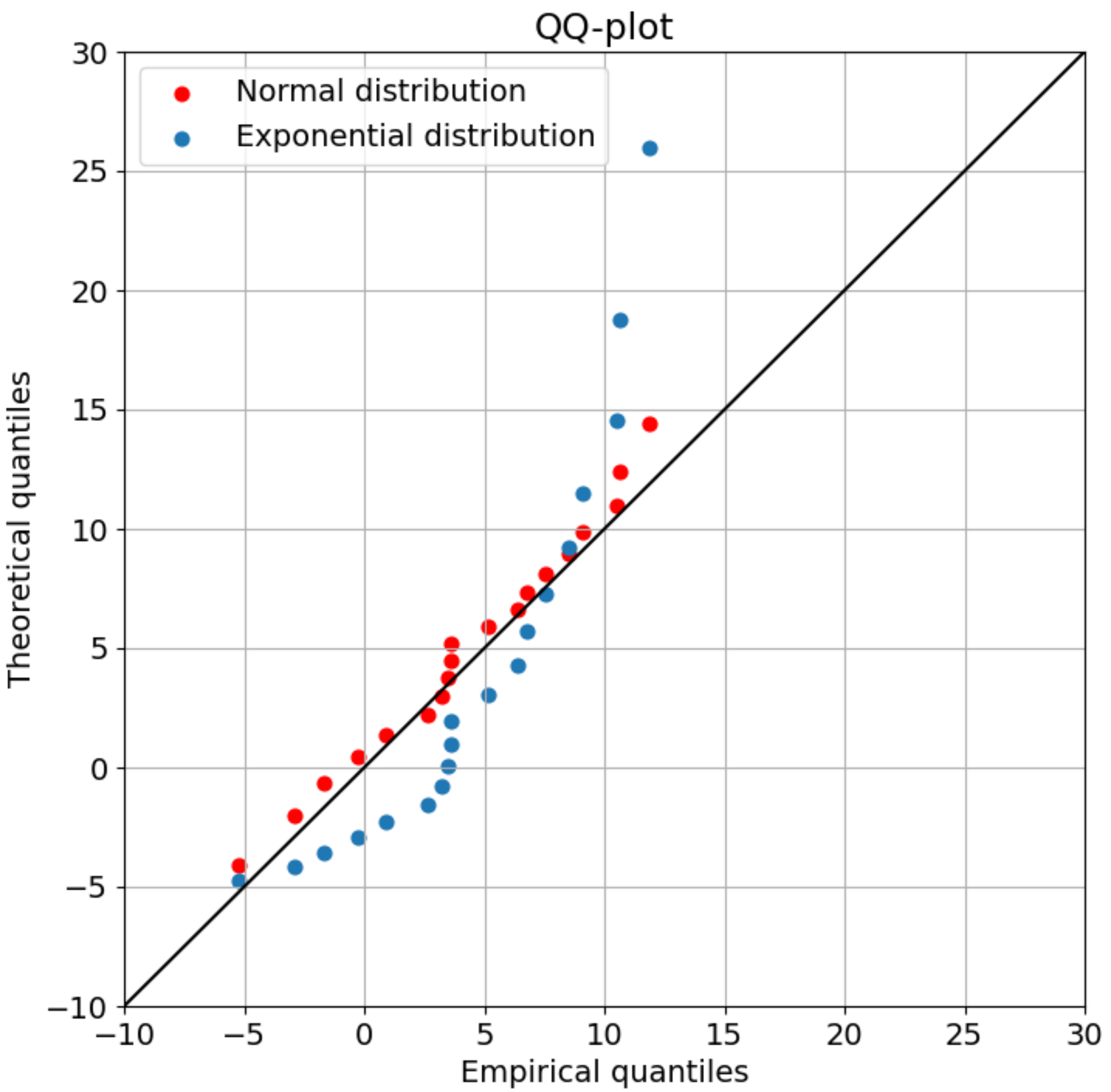
TU Delft

QQ-plot

- X-axis: quantiles of the observations
  $X = \{19.54; 9.12; 11.89; -0.29; 2.65; 3.63; 10.49; \ldots$
  $3.61; 8.50; -5.25; 3.23; 0.88; -2.88; 7.53; 6.40; \ldots$
  $5.16: -1.66; 10.63; 6.75; 3.50\}$

- Y-axis: quantiles predicted by the fitted distribution, $N(5.17, 5.76)$ or Exp(-5.25, 0.10)

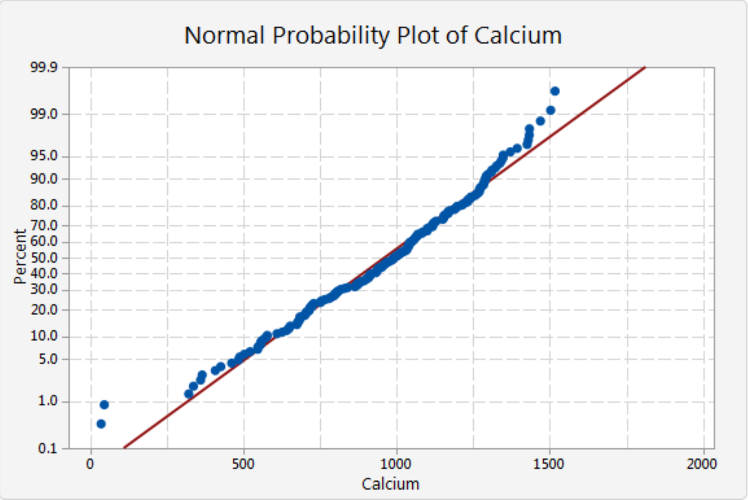
- Perfect fit: 45-degrees line



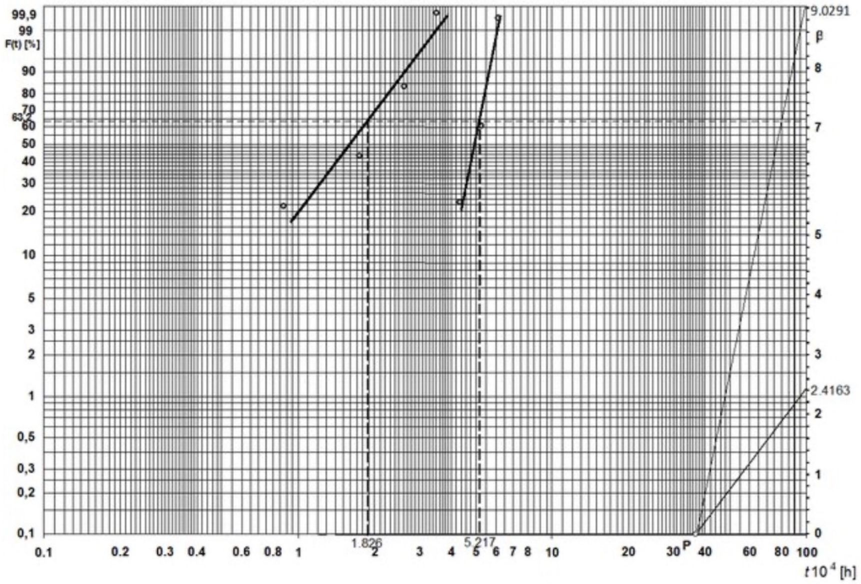
TU Delft

Probability plot

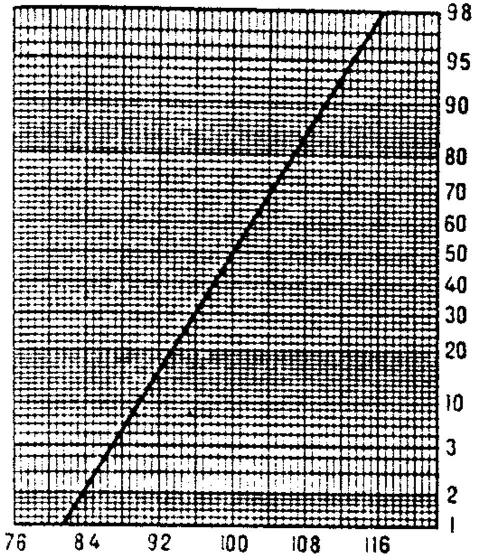
- Grid of one axis is adapted to a theoretical distribution function, so when it is plotted, a line is obtained.



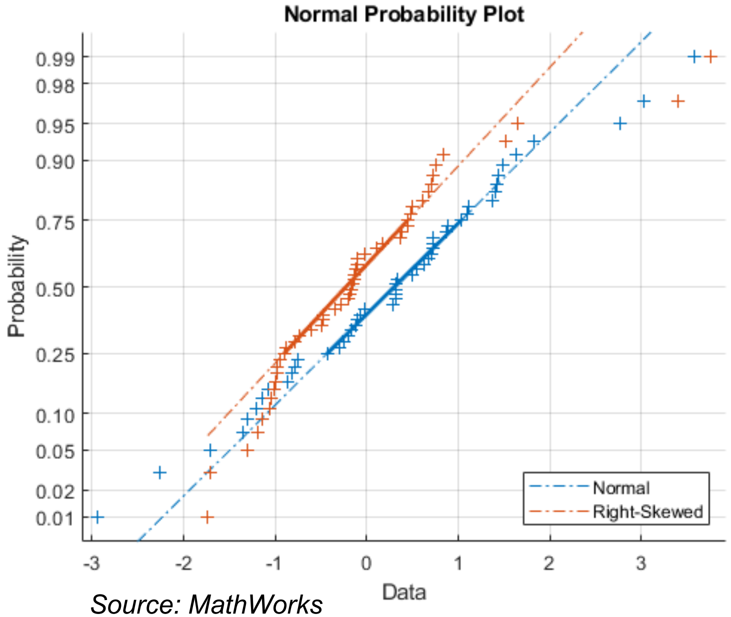
Source: https://stats.stackexchange.com/questions/554193/what-determines-y-axis-scaling-on-a-normal-probability-plot



Source: Kalaba et al. (2014)



Source: https://encyclopediaofmath.org/wiki/Probability_graph_paper



Source: MathWorks

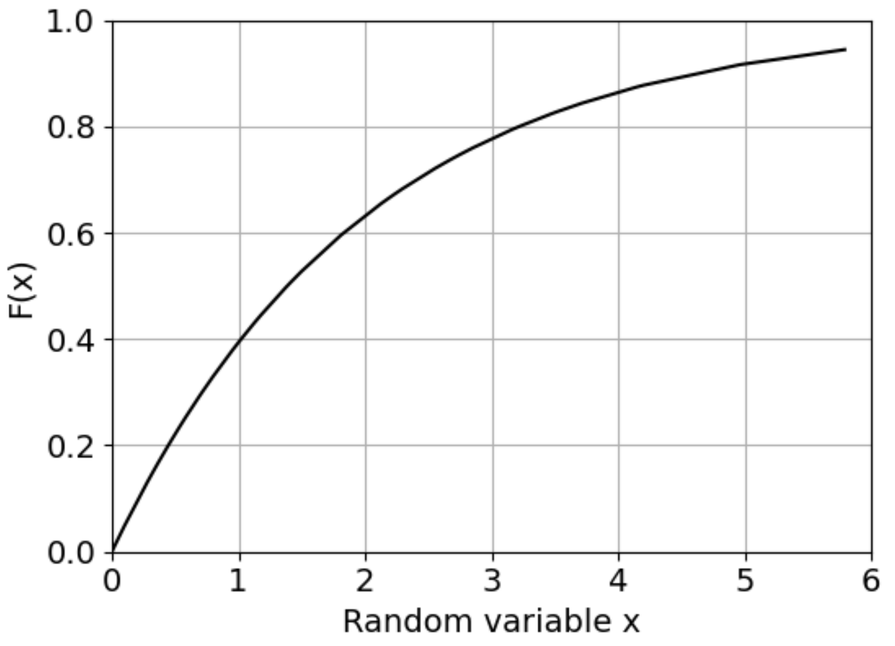
Probability plot

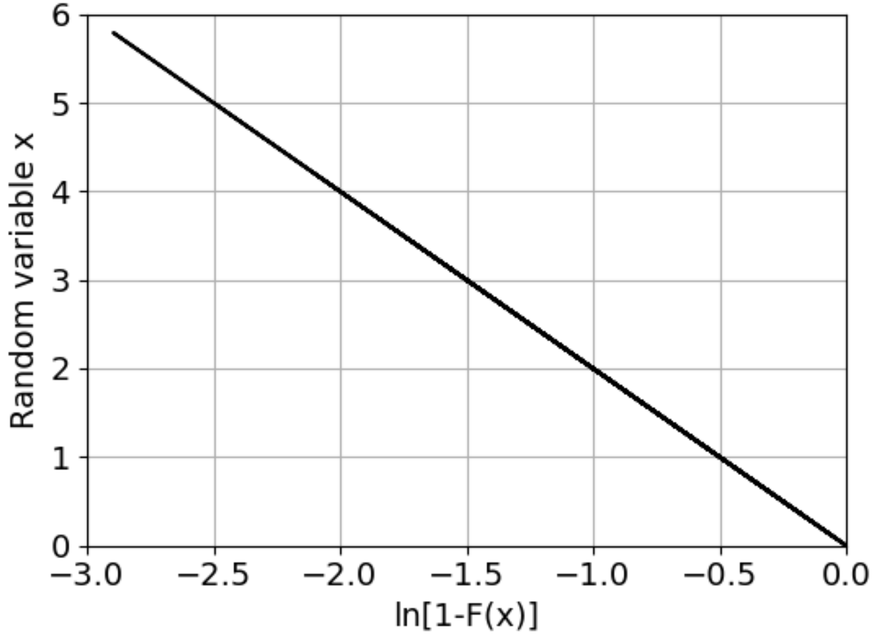
- Grid of one axis is adapted to a theoretical distribution function, so when it is plotted, a line is obtained.

- <u>Example:</u> Exponential cdf

$$F(x)=1-\exp(-\lambda(x\text{-}\mu)) \rightarrow \ln[1-F(x)]=-\lambda(x\text{-}\mu)$$

- $\ln[1-F(x)]$ vs. x → Linear relationship





TUDelft

Probability plot (example)

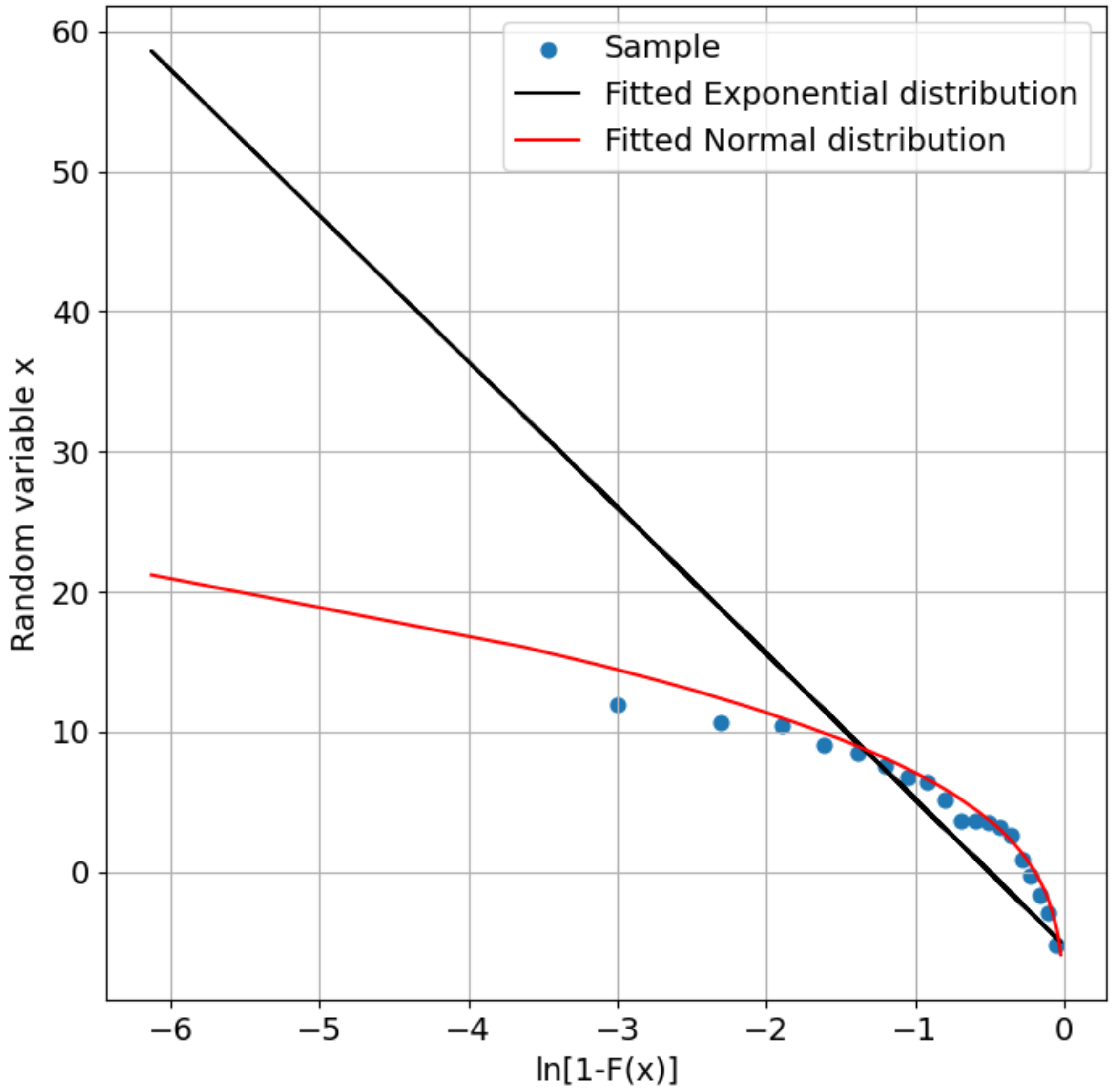
- $X = \{19.54;9.12;11.89;-0.29;2.65;3.63;10.49;3.61;8.50;-5.25;\ldots$
  $3.23;0.88;-2.88;7.53;6.40;5.16;-1.66;10.63;6.75;3.50\}$

- Fitted $N(5.17,5.76)$

- Fitted Exp(-5.25, 0.10)

- X-axis: ln[1−$F(x)$]

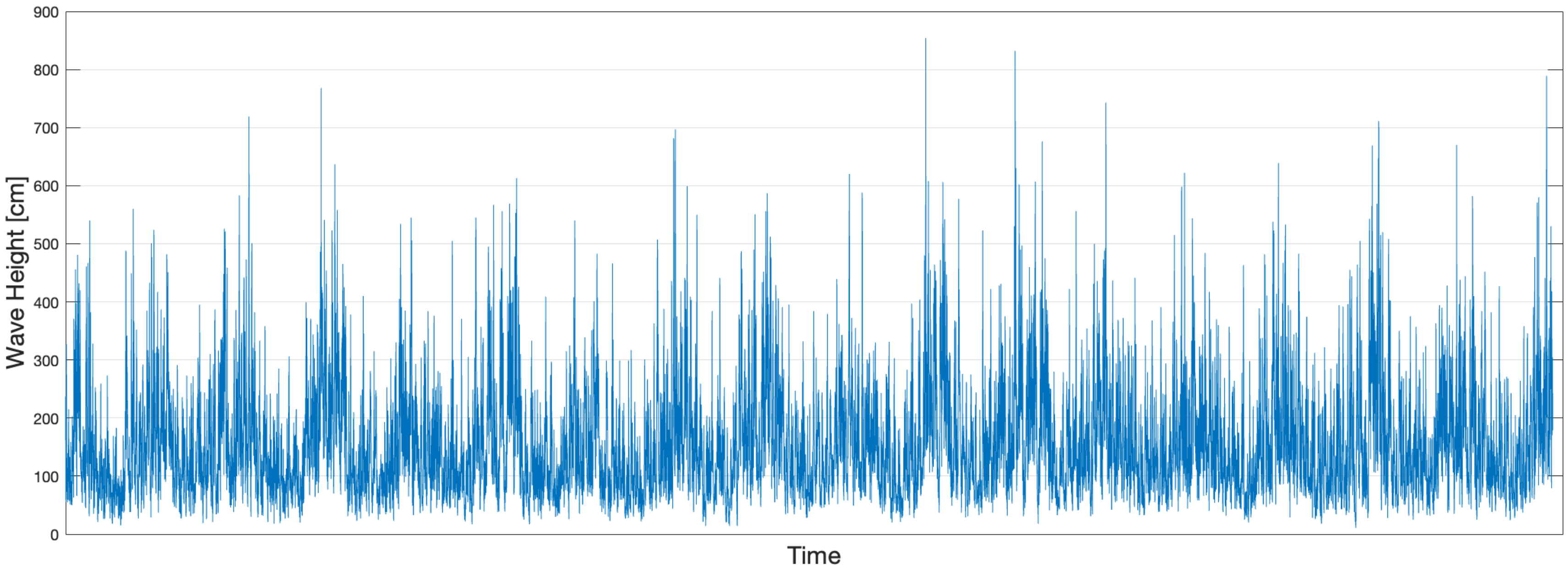
- Y-axis: x



TU Delft

Let’s get practical

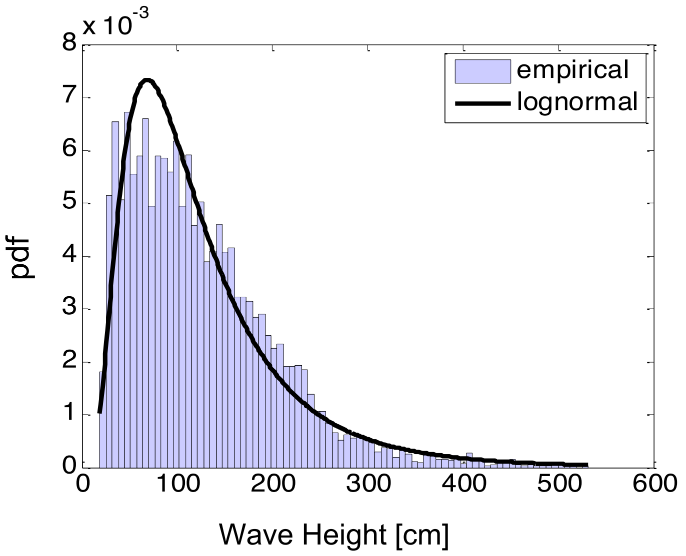
- Observations wave height [cm] Europlatform

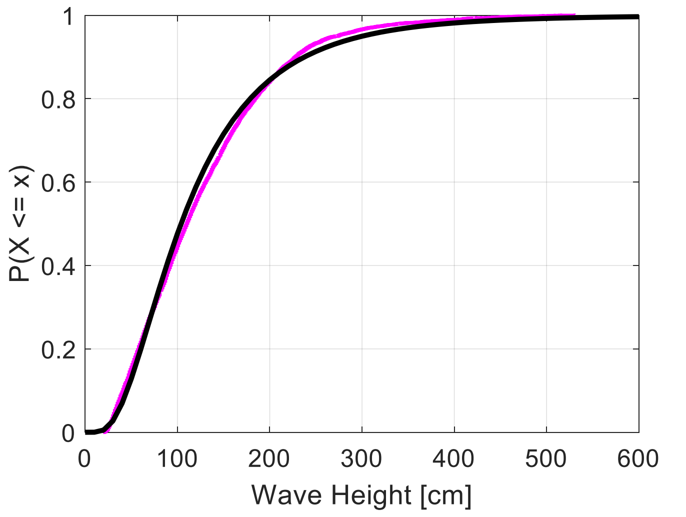


TU Delft

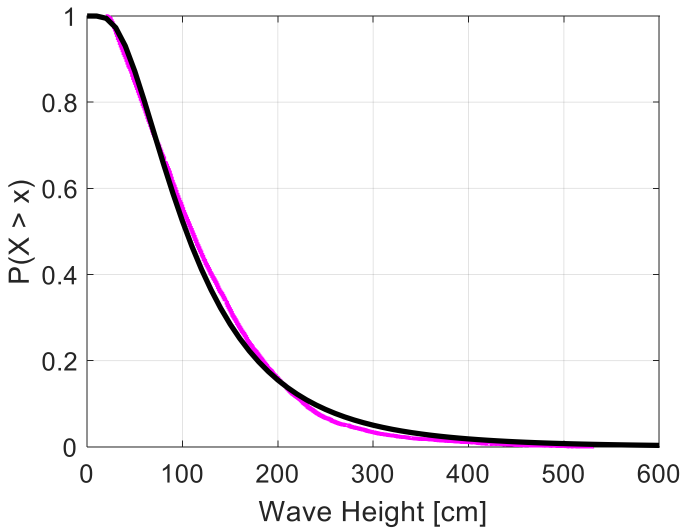
Is it a good model?

- Maximum likelihood fit to a Lognormal distr.
    - mean = 127.9,
    - sd = 91.91
    - pdf and cdf





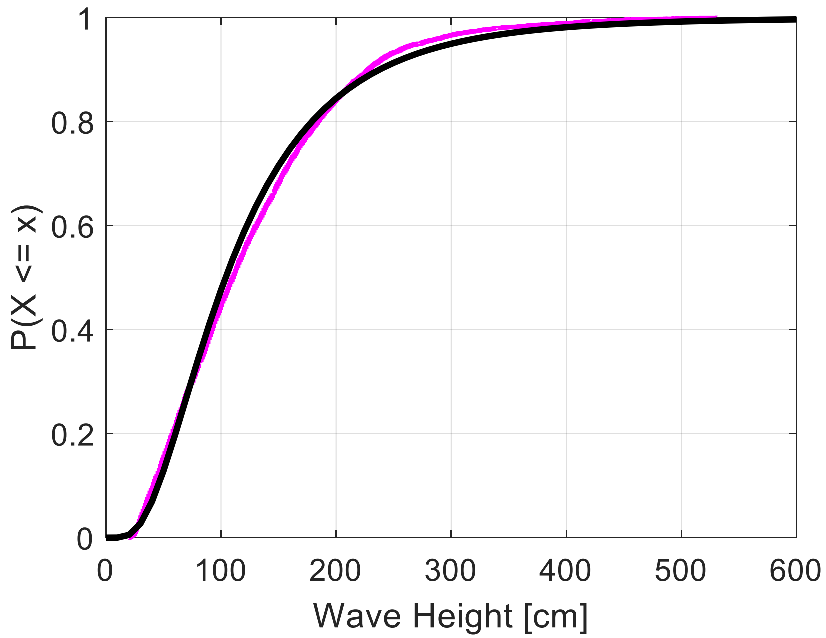
- $P(X>x)=1-F_x(x)$
- Sometimes called
    - Survivor or Reliability function
    - Exceedance probability

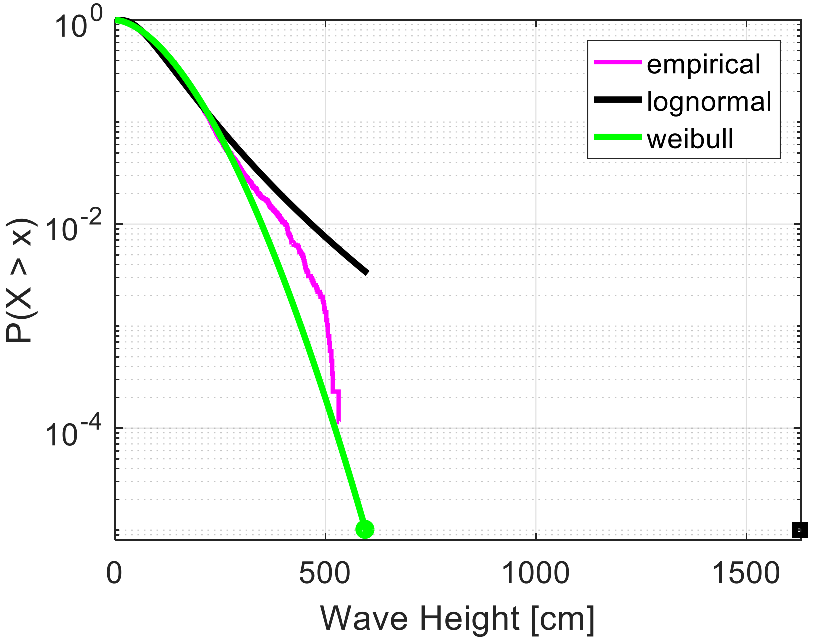


Is it a good model?

- For design purposes
  - Find x_{large} s.t. $P(X > x_{large}) = 10^{-5}$

- Extrapolate through parametric distribution
  - Lognormal x_{large}=15.30m
  - Weibull x_{large}=5.95m

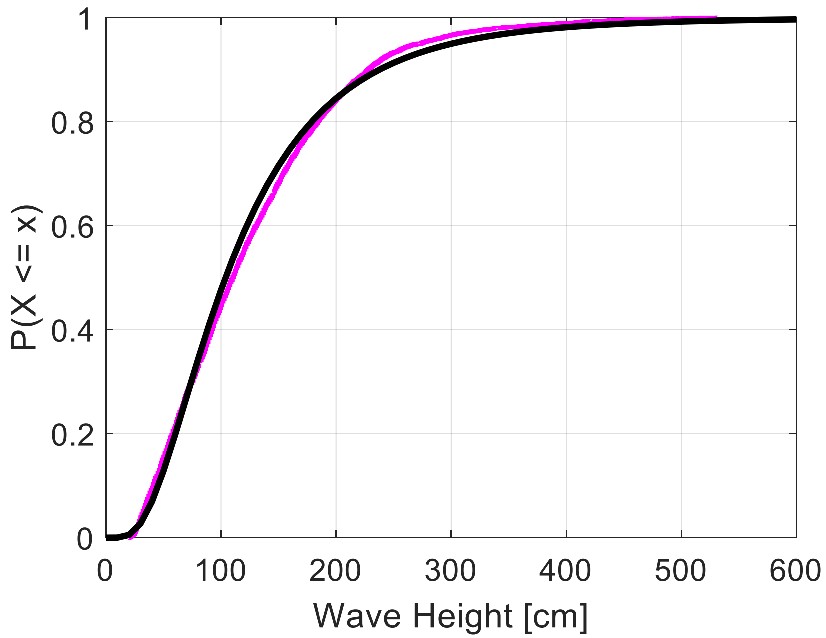


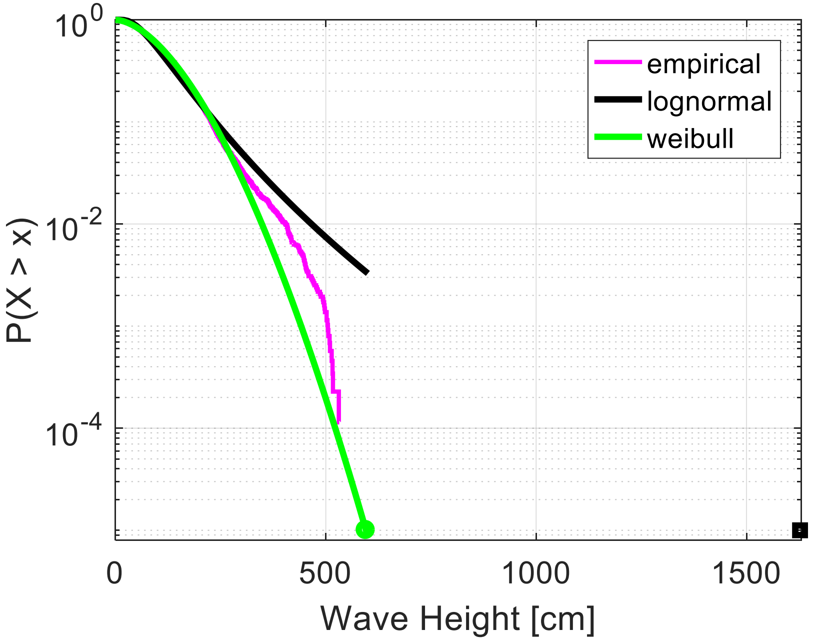


TU Delft

Use common sense!

- Be careful with extrapolation
- “Very large” or “very small” values (tails of the distribution) are extreme values
  - Extreme Value Analysis

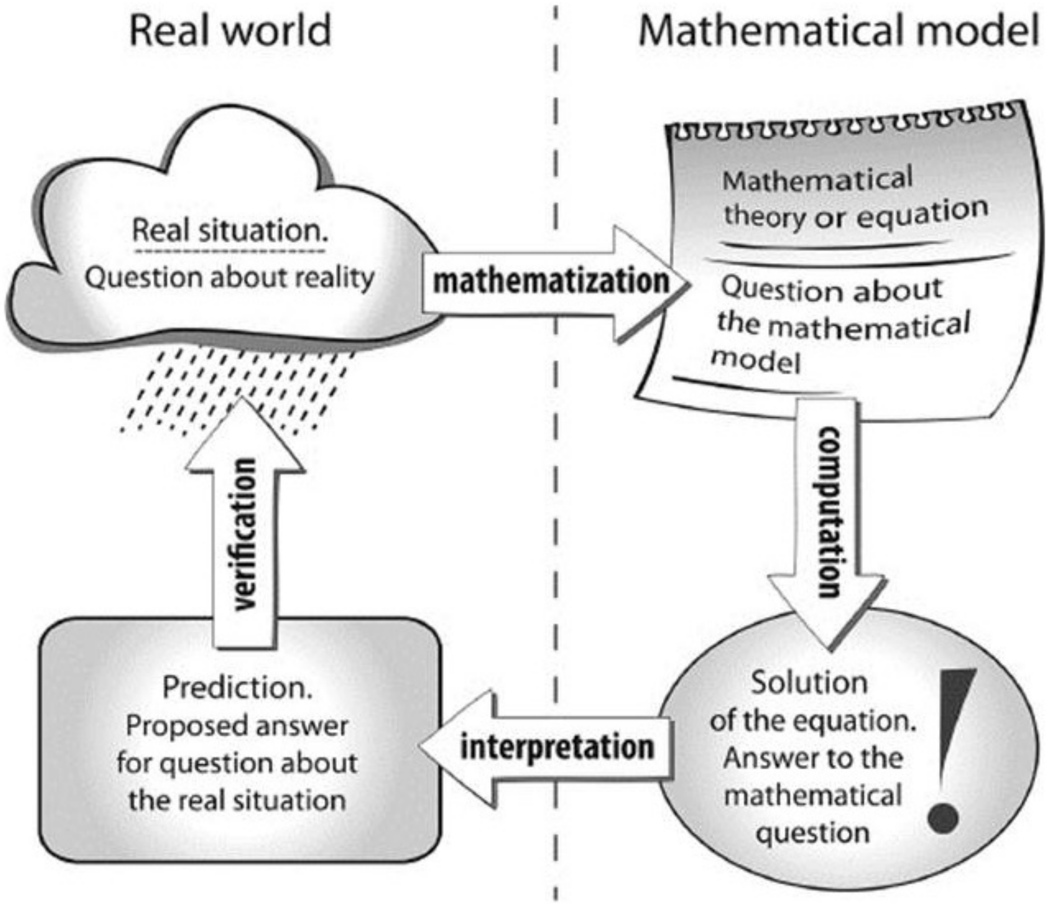




TU Delft

Final Remarks

- Good models are those that are useful.
- No model is perfect. If it were, it would not be a model.
- Best check is common sense. A good model is one which provides reasonable useful answers.



Source: https://schoolbag.info/mathematics/numbers/103.html