

1 Einleitung

1.1 Das Grundschemata der parametr. Statistik

a Wahrscheinlichkeits-Theorie: **Modell**.

Typischerweise parametrische Familie, z.B. Normalvert. $\mathcal{N}(\mu, \sigma^2)$.

Statistik: Brücke zwischen Modell und Daten. **Drei Grundfragen der Schliessenden Statistik**:

[1.] Welcher Wert ist für den Parameter **am plausibelsten**?

→ **Schätzung**

[2.] Ist ein bestimmter Wert plausibel?

→ **Test**.

[3.] Welche Werte sind insgesamt plausibel?

→ **Vertrauens- oder Konfidenzintervall**

b **Wahrscheinlichkeitsmodell** wird gebraucht, um zu beschreiben, was „auch noch hätte herauskommen können, und mit welchen Chancen“.

Besser: W.modell besteht, **bevor wir die Daten sehen**, und beschreibt unsere Vorstellung, was für Resultate wir mit welcher „Plausibilität“ erwarten.

Grundlage für die Bestimmung der statistischen Streuung einer Schätzung oder Test-Statistik.

c Voraussetzungen überprüfen!

- d Am liebsten hätte man Methoden, für die **Abweichungen von den Voraussetzungen überhaupt keine Rolle** spielen.
- e Rangsummentests sind von Verteilungsannahme (weitgehend) unabh.
- f **Grundidee des Resampling:**
Die Daten selber verwenden, um ihre Verteilung zu schätzen.
- g Komplizierte Schätzung \rightarrow Genauigkeit über
 - Bootstrap, ...
 - Asymptotische Verteilung.

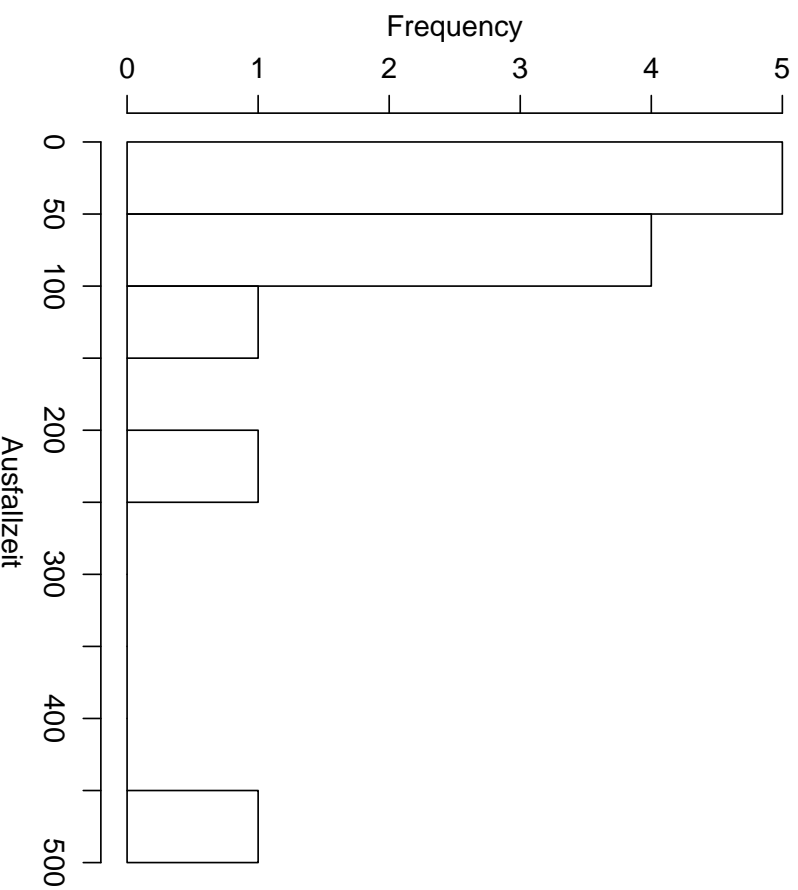
1.2 Beispiel

a **Ausfallzeiten** des Air conditioning-Systems in Boeing 720

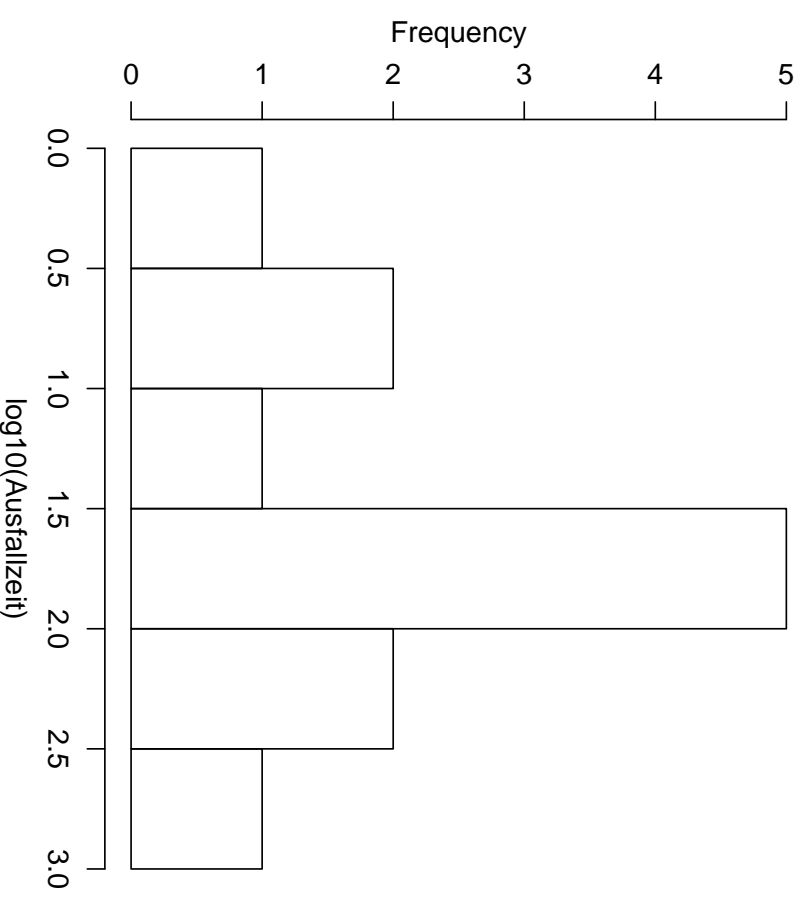
$n = 12$ Zeiten zwischen Ausfällen (sortiert):

3 5 7 18 43 85 91 98 100 130 230 487

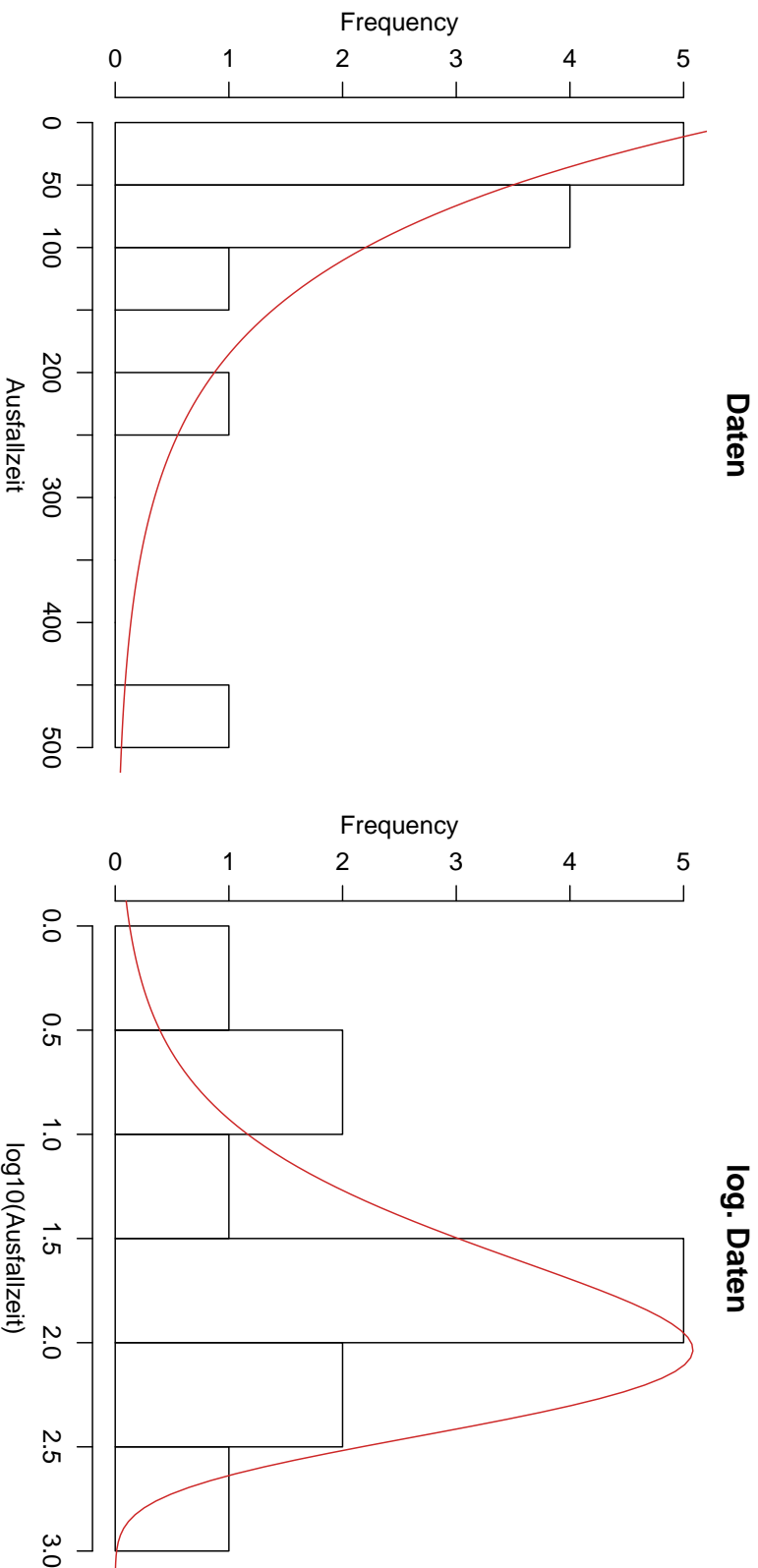
Daten



log. Daten



- b Einfachstes parametrisches Modell für Ausfallzeiten:
Exponential-Verteilung $\&xp$ mit Dichte



c Modell passt nicht schlecht. Aber der Datensatz ist klein.

Besser: **Keine Verteilung voraussetzen!**

Frage: Mittelwert? $\bar{x} = 108.1$.

Etwas spannender: 20% gestutztes Mittel?

= Lasse die 20% kleinsten & 20% grössten Daten weg,
bilde Mittel der übrigen!

Für $n = 12$ je 2 Beob. weglassen.

→ $(7 + 18 + 43 + 85 + 91 + 98 + 100 + 130) / 8 = 71.5$

„Eine Zahl ohne Genauigkeitsangabe ist wertlos!“

→ Schliessende Statistik, Vertrauensintervall.

Dafür braucht man Wahrscheinlichkeitsmodelle!