

# Design of Computer Experiments and Active Learning

Seminar

JProf. Dr. Kirsten Schorning

Mathematische Statistik  
Statistik  
TU Dortmund

# Topics and Research Areas in Statistics



# Does the Design of Experiments matter?

---

**Simple Linear Regression Model:** The relationship between an input  $x$  and an output  $y$  is modelled by

$$Y = ax + b + \varepsilon,$$

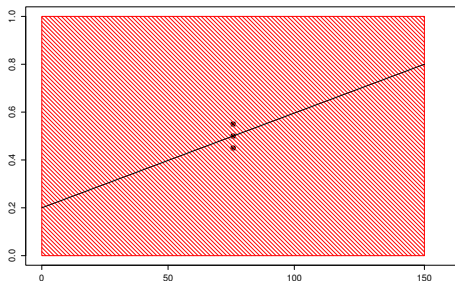
where  $(a, b)^T \in \mathbb{R}^2$  are unknown.

**Question:** Which input positions should be selected to get informative observations that result in a precise fit?

# Does the Design of Experiments matter? Example

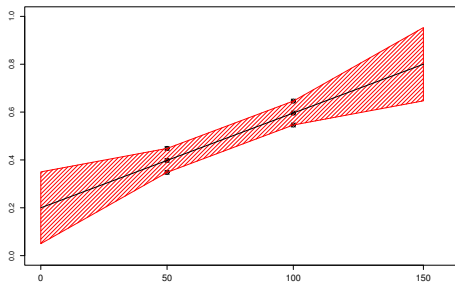
---

1. **Idea:** All observations are taken at one point, here at 75.



# Does the Design of Experiments matter? Example

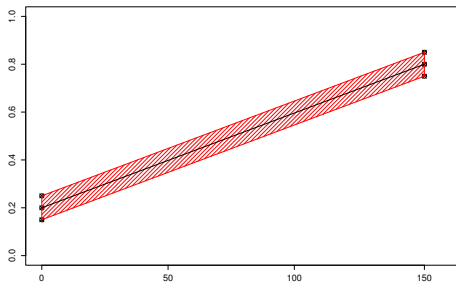
2. **Idea:** Half of the outputs are measured at 50, half of the outputs are measured at 100.



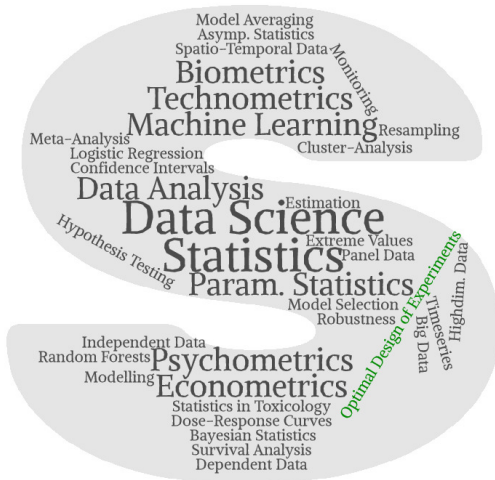
# Does the Design of Experiments matter? Example

---

2. **Idea:** Half of the outputs are measured at 0, half of the outputs are measured at 100.



# The Design of Experiments matters!



# The content of the seminar I

---

- Optimal Design for Computer Experiments
  - ▶ Target: Designing Computer Experiments such that Predictions based on the computer experiments become precise.
  - ▶ Topics: The Stochastic Process Model, Space-filling Designs, Latin Hypercube Designs, Criterion-based experimental designs
  - ▶ Literature:
    - ★ Santner, Williams, Notz (2003). The Design and Analysis of Computer Experiments. Springer, New York. see link for E-Book
    - ★ Further papers on the performance of certain design strategies.



# The content of the seminar II

---

- Active Learning for Machine Learning

- ▶ Target: Improve statistical learning algorithms by letting them take an active role in the selection of which examples are labeled, e.g. reach a desired error rate using much fewer labels than needed if random sampling is done.
- ▶ Topics: Agnostic Active Learning, Disagreement-based active learning, cluster-based active learning.
- ▶ Literature:
  - ★ Settler, Burr (2012). Active Learning. Springer Cham, Switzerland. See link for E-book
  - ★ Further papers on active learning, e.g. on agnostic active learning.

# Time Schedule

---

Registration (until **February, 22nd, 2023**):

- Please have a look in the two E-books before deciding for the seminar!
- E-Mail to [schorning@statistik.tu-dortmund.de](mailto:schorning@statistik.tu-dortmund.de) and state whether you prefer to do a presentation on DOE in Computer Experiments or on Active Learning.

Kick-off meeting (including the assignment of seminar topics to participants)

- Tuesday, 07.03.2023, 10 am.

Seminar presentations:

- Each Monday, 2.15-3.45 pm starting April, 17th, in M/E27.
- Some extra seminar meetings if needed (depends on the number of participants).

Please sign in for the moodle course: links follows as soon as possible.

# Requirements to pass the seminar

---

## **CAUTION**

- The registration and meeting are obligatory, if you want to participate in the seminar.

## **HOW TO PASS**

- presentation in English or German with English slides (45 minutes)
- seminar paper either English or German (10 pages)
- active participation in discussions, feedback

## **FEEDBACK ROUND**

- The slides of the presentation have to be sent to me one week before the presentation will take place.
- After my review, you will have to revise the slides.