Introduction

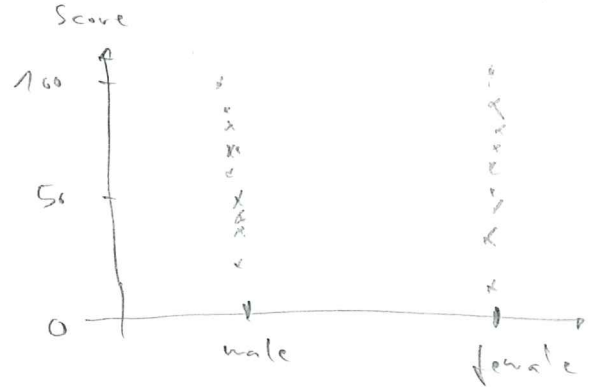
Real-life example

1. Alien invasion from Mars (Mars Attacks, 1996)

   a) Investigate: different humans: male / female

   b) Experiment: BBQ and taste the humans

      Assign taste score (0 - 100)
      bad ... good

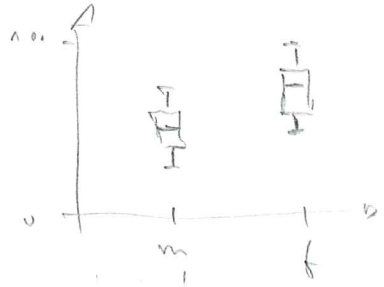


   c) Question: who tastes better, males or females?

Aliens are stupid ⇒ How to analyze this? Your turn!

Response:

- Visualization: box plot

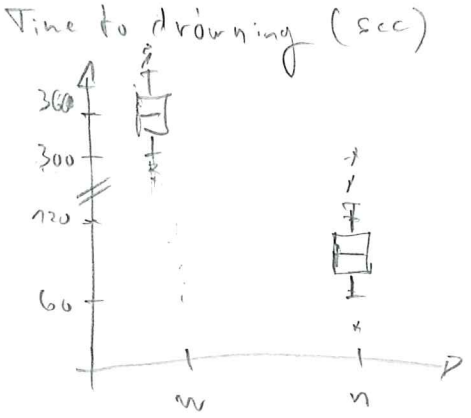


- Hypothesis testing: e.g. t-test, $H_0: \mu_m = \mu_f$ (mean male, mean female)

2. Abduction by Nazi scientists

a) Question: how good is the new "wetsuite-uniform" for pilots

b) Experiment: throw you in ice water, either with wetsuite or standard uniform



Nazis also stupid

$\Rightarrow$ How to analyze this setting, what is different to the previous one?

Response:

- Experimental hypothesis: $\mu_w > \mu_n$
- Leads to so-called "one-tailed" test, rare in practice

Historical note: Sad, but similar experiments were really carried out!

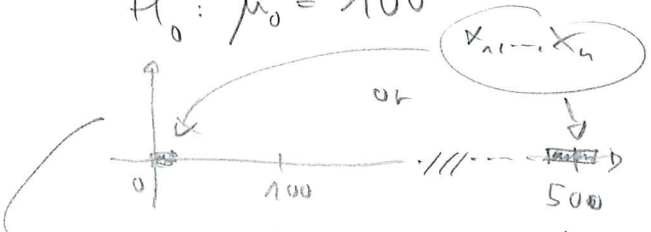
Recap - testing principles

- Situation: data, experimental hypothesis, other info

$x_1, \dots, x_n$: "pill should do something", "$\mu = 100$ without pill", "$\sigma_x = 15$", ...

- Idea: formulate statistical hypoth., reject it if observations are "too extreme"

$H_0: \mu_0 = 100$ (or simply μ)



→ in principle good idea, but:
- interval range(s)?
- what to do for next experiment?

- Improve: Look at appropriate test statistic for the setting

$x_1, \dots, x_n$ Gaussian, σ_x known

$$\Rightarrow \frac{\bar{x} - \mu_0}{\frac{\sigma_x}{\sqrt{n}}} \sim N(0,1) \text{ if } H_0 \text{ is true}$$

from data, known facts, H_0

works for all similar experiments

- Result: standardized test, works in all similar settings

z-test, based on standard Gaussian