

## Phd-kursus i Basal Statistik, Opgaver til 2. uge

### Opgave 1: Sædkvalitet

Datafilen `oeko`, som du finder i forskellige formater på hjemmesiden med datasæt, indeholder datamateriale til belysning af forskellen i sædkvalitet mellem SAS-ansatte og mænd, der lever økologisk. Variablene er (i den nævnte rækkefølge):

- `sas_ansat`: ansat i SAS (ja/nej)
- `abstid`: abstinensid (1: kort, 2: medium, 3: lang)  
(et mål for længden af seksuel afholdenhed)
- `konc`: sædkoncentrationen (mill/ml)

**Formålet med opgaven er at undersøge, om der er forskel på de to populationsgruppers sædkoncentrationsniveau.**

Husk undervejs at overveje kontrol af modelantagelserne, uden at lade dette tage fokus fra selve analyserne og fortolkningerne.

1. Lav en passende illustration af data.
2. Vi skal nu kvantificere niveauet af sædkoncentration for de to grupper af mænd og sammenligne disse niveauer, i første omgang uden at tage hensyn til abstinensiden. Overvej, om der skal logaritmetransformeres, når I svarer på nedenstående spørgsmål:
  - (a) Giv et estimat for niveauet af sædkoncentrationen for hver af de to grupper af mænd. Husk et 95% konfidensinterval.
  - (b) Sammenlign de to estimater og de to tilhørende konfidensintervaller fundet ovenfor, og giv en intuitiv vurdering af, hvorvidt der er forskel på de to grupper eller ej.
  - (c) Foretag nu en sammenligning af de to grupper, og kvantificer forskellen i sædkoncentration for grupperne, igen med 95% konfidensinterval.
  - (d) Er der signifikant forskel på de to gruppers sædkoncentration?

3. Lav nu en underopdeling af de to grupper, baseret på længden af abstinenstiden og udregn passende summary statistics for de nu i alt 6 grupper.
  - (a) Ser det ud som om abstinenstiden har indflydelse på sædkoncentrationen?  
Og i givet fald, ser denne indflydelse så ens ud i grupperne?
  - (b) Ser det ud som om fordelingen af abstinenstider er den samme i de to grupper?
4. Benyt en variansanalysemodel til at besvare følgende:
  - (a) Find et estimat for forskellen i sædkoncentration mellem de to populationer af mænd, for fastholdt værdi af abstinenstid.
  - (b) Sammenhold ovenstående estimat med det tilsvarende fra spørgsmål 2c og kommenter:  
Kan man sige, at forskelle i sædkoncentration i de to grupper skyldes forskelle i deres abstinenstider?
  - (c) Er der evidens for, at abstinenstiden har en forskellig effekt på sædkoncentrationen i de to populationer?

## Opgave 2: Space shuttle (9.1, p. 223)

Et studie involverer de 26 astronauter, der deltog på de første 8 rejser med space shuttle (Bungo et.al., 1985).

På frivillig basis konsumerede 17 af disse astronauter store mængder af salt og væske inden landingen, i et forsøg på at imødegå 'space deconditioning' (salt=1).

De 9 øvrige indtog intet specielt (salt=0). Tabellen nederst viser pulsen (slag pr. minut) før og efter flyvningen for hver af de 26 astronauter.

Data er lagt i filen `space` (findes i forskellige formater på hjemmesiden med datasæt), med 3 kolonner af tal og med variabelnavnene `salt`, `pre` og `post` i første linie.

1. Sammenlign før- og efter-målingerne i 'salt'-gruppen, og husk samtidig at lave passende illustrationer.

2. Huskede du at give en kvantificering af effekten af flyvning på pulsen i 'salt'-gruppen? Med konfidensinterval!
3. Sammenlign effekten af flyvning i de to grupper. Hvilken konklusion opnås for effekten af saltindtagelse? Husk konfidensinterval!
4. To astronauter deltog i to forskellige flyvninger og optræder altså i datamaterialet to gange. Spiller det nogen rolle?
5. Kommenter frivilligheden i opdelingen i de to grupper og hvordan dette kunne tænkes at påvirke fortolkningen af resultaterne.

| Countermeasure taken |       |        | Countermeasure not taken |       |        |
|----------------------|-------|--------|--------------------------|-------|--------|
| Pre                  | Post  | Change | Pre                      | Post  | Change |
| 71                   | 61    | -10    | 61                       | 61    | 0      |
| 65                   | 59    | -6     | 59                       | 66    | 7      |
| 52                   | 47    | -5     | 52                       | 61    | 9      |
| 68                   | 65    | -3     | 54                       | 68    | 14     |
| 69                   | 69    | 0      | 53                       | 77    | 24     |
| 49                   | 50    | 1      | 78                       | 103   | 25     |
| 49                   | 51    | 2      | 52                       | 77    | 25     |
| 57                   | 60    | 3      | 54                       | 80    | 26     |
| 51                   | 57    | 6      | 52                       | 79    | 27     |
| 55                   | 64    | 9      |                          |       |        |
| 58                   | 67    | 9      |                          |       |        |
| 57                   | 69    | 12     |                          |       |        |
| 59                   | 72    | 13     |                          |       |        |
| 53                   | 69    | 16     |                          |       |        |
| 53                   | 72    | 19     |                          |       |        |
| 53                   | 75    | 22     |                          |       |        |
| 48                   | 77    | 29     |                          |       |        |
| Mean                 | 56.88 | 63.76  | 57.22                    | 74.67 | 17.44  |
| SD                   | 7.30  | 8.86   | 8.44                     | 13.01 | 10.11  |