

Afleveringsopgave 1: Ensidet variansanalyse med forskellig gruppestørrelse

Opgaven afleveres i papirform til Anders Tolver eller en af instruktorerne.

Aflevering: Senest tirsdag d. 12. september 2017 kl. 10:00.

Opgaven skal benyttes til både at indarbejde brugen af R og til at beskrive den statistiske analyse og fortolkningerne af resultaterne fra den.

Hvis I arbejder sammen i en gruppe afleveres kun en besvarelse per gruppe.

På en virksomhed, hvor man fremstiller gødningsstoffer, har der været fremført klager over et varierende kali-indhold (K_2O) fra fire forskellige produktionsgrene. For at undersøge sagen har man foretaget en række målinger, som findes i datasættet `Afl1data.txt`¹. Bemærk at der ikke er lige mange målinger for hver produktionsgren.

På baggrund af datasættet ønskes følgende 7 spørgsmål besvaret. Besvarelsen må meget gerne beskrive, hvordan R er benyttet ved opgaveløsningen. En oplagt mulighed er at lave et R markdown dokument, hvori man indfører kommentarer, der (kort) forklarer, hvad de vigtigste dele af koden gør. Selve opgavebesvarelsen kan så bestå af et dokument der fås frem ved at knytte dit R markdown dokument fx. til et word-dokument. (Det er en god vane altid at skrive udførlige kommentarer til både R-kode og output, som så kan gemmes og tages frem senere!)

1. Lav et boxplot, der viser indholdet af K_2O for hver produktionsgren.
2. Opstil modellen for ensidet variansanalyse.
3. Angiv parameterestimerne for indholdet af K_2O i hver produktionsgren samt estimeret for spredningen ($s = \hat{\sigma}$).
4. Test hypotesen om, at der ikke er forskel på K_2O -indholdet i de 4 produktionsgrene. Angiv resultaterne i et modelskema og et testskema (-søg inspiration i slides fra forelæserne i uge 1).
5. Angiv 95 %-konfidensintervaller for parameterestimer under slutmodellen.
6. Angiv LSD værdien for sammenligning mellem K_2O -indholdet i gren A og B.
7. Er der forskel på behandlingerne givet ved produktionsgren C og D.
8. Lidt mere udfordrende: Giv et forslag til, hvordan man kunne teste, om det i virkeligheden kunne forholde sig sådan, at indholdet i grenene B, C og D er ens, og at gren A adskiller sig fra de øvrige tre grene.

¹Data stammer fra eksempel 5.1 i Knut Conradsen: En introduktion til statistik, IMM