Første sett med obligatoriske oppgaver
STK2120 våren 2016

Her er det første settet med obligatoriske oppgaver i STK2120 våren 2016. Oppgavesettet består av tre oppgaver. Prøv å besvare spørsmålene kort og konsist, men likevel med gode forklaringer! Der du bruker R, må kommandoer og resultater innarbeides i rapporten eller legges ved.

Oppgavene er obligatoriske og studenter som ikke får besvarelsen godkjent, vil ikke få adgang til avsluttende eksamen. For å få besvarelsen godkjent, må du minst ha gjort et forsøk på å løse alle deloppgaver.

Det er helt i orden og utmerket om du samarbeider med andre og diskuterer hvordan oppgavene skal løses. Den innleverte besvarelsen skal imidlertid være skrevet av deg og gjenspeile din forståelse av stoffet. Det må gå fram av besvarelsen hvem du eventuelt har samarbeidet med. Er vi i tvil om du har forstått det du har levert inn, kan vi be deg om en muntlig redegjørelse.

Besvarelsen leveres ved Matematisk institutt, 7. etasje, Niels Henrik Abels hus.
Husk at du skal bruke Matematisk institutts forside ved innlevering. Du finner denne her:
`www.uio.no/studier/admin/obligatoriske-aktiviteter/mn-math-obligforside.pdf`

Frist for innlevering er torsdag 18. februar kl. 14.30.

Oppgave 1

Kudzu er en plante som inneholder isoflavoner. I en studie ville en undersøke om isoflavoner kan ha en gunstig effekt for skjelettet ved at de øker beintettheten. 45 rotter ble delt tilfeldig inn i tre grupper med 15 rotter i hver gruppe. En gruppe fikk en lav dose av isoflavoner fra kudzuplanten i maten, en gruppe fikk en høy dose og en kontrollgruppe fikk ikke noe tilskudd av isoflavoner. Etter en tid målte en beintettheten i lårbeinet til rottene. Resultatet er gitt i tabellen nedenfor.

Behandling	Beintetthet (g/cm^2)							
Kontroll	0.228	0.207	0.234	0.220	0.217	0.228	0.209	0.221
	0.204	0.220	0.203	0.219	0.218	0.245	0.210	
Lav dose	0.211	0.220	0.211	0.233	0.219	0.233	0.226	0.228
	0.216	0.225	0.200	0.208	0.198	0.208	0.203	
Høy dose	0.250	0.237	0.217	0.206	0.247	0.228	0.245	0.232
	0.267	0.261	0.221	0.219	0.232	0.209	0.255	

Du kan lese dataene inn i R ved kommandoen:

```
bonedensity=
read.table("http://www.uio.no/studier/emner/matnat/math/STK2120/v16/bonedensity.txt",
header=T)
```

I den datarammen du får, er behandlingen gitt i første kolonne (1 = kontroll, 2 = lav dose, 3 = høy dose) og beintettheten gitt i andre kolonne.

Vi er interessert i å undersøke om isoflavoner har betydning for beintettheten.

a) Vi tenker oss at beintettheten for rotte j ved behandling i er en observert verdi av den stokastiske variabelen X_{ij}. Formuler en rimelig model for X_{ij} - ene.

b) Ta utgangspunkt i modellen fra punkt a og forklar hvordan du kan bruke variansanalyse for å undersøke om isoflavoner har betydning for beintettheten. Gjennomfør variansanalysen fortolk resultatene av den.

Oppgave 2

For rullestolbrukere kan trykkbelastninger være et problem som i verste fall kan føre til sårdannelser i setemusklene. Spesielle rullestolputer kan redusere trykkbelastningen. Vi skal i denne oppgaven se nærmere på en studie som sammenliknet fem ulike typer rullestolputer: Reston Floatation Pad (RF), Stryker Flotation Pad (SF), Spenco Skin Care Pad (SC), Bye-Bye Decubiti (BD) og Jobst Hydro-Float Pad (JP).

Ti frivillige voksne personer var med i studien. Hver av personene prøvde de fem putetypene i tilfeldig rekkefølge, og trykket (i mmHg) mellom setemusklene og rullestolen ble målt. Resultatet er gitt i tabellen nedenfor.

Person nr.	Pute RF	Pute SF	Pute SC	Pute BD	Pute JP
1	47	59	40	49	40
2	43	47	45	49	36
3	42	43	38	30	41
4	76	84	64	46	49
5	49	70	57	58	29
6	62	86	55	63	67
7	53	68	51	50	48
8	75	61	58	56	34
9	88	80	76	82	62
10	77	76	88	56	75

Du kan lese dataene inn i R ved kommandoen:

```
rullestol=
read.table("http://www.uio.no/studier/emner/matnat/math/STK2120/v16/rullestol.txt",
header=T)
```

I den datarammen du får, er nummeret på personen gitt i første kolonne, type rullestolpute i andre kolonne og trykket i tredje kolonne.

a) Studien av rullestolputene ble gjort som et blokkforsøk, der hver person prøvde alle de fem putene. Kan du beskrive en annen type forsøk som kunne ha vært benyttet for å sammenligne putene? Hvorfor tror du det ble valgt å gjøre et blokkforsøk?

b) Vi tenker oss at trykkmålingen for pute i og person j er en observert verdi av den stokastiske variabelen X_{ij}. Formuler en rimelig model for X_{ij} - ene.

c) Ta utgangspunkt i modellen fra punkt b og forklar hvordan du kan bruke variansanalyse for å undersøke om det er forskjell på putene. Gjennomfør variansanalysen fortolk resultatene av den.

d) Forklar hvordan du kan bruke Tukeys metode for å bestemme simultane konfidensintervaller for de forventede forskjellene i trykkbelastning mellom putene. Beregn de simultane konfidensintervallene med simultan konfidenskoeffisient 95% og forklar hva de sier deg om forskjellen mellom putene.

Oppgave 3

I denne oppgaven vil vi se nærmere på modellen for enveis variansanalyse når det er like mange observasjoner i hver gruppe. Modellen kan gi som

$$X_{ij} = \mu + \alpha_i + \epsilon_{ij}\,; \quad j = 1, 2, \ldots, J\,; \quad i = 1, 2, \ldots, I\,; \tag{1}$$

der ϵ_{ij} - ene er uavhengige og $N(0, \sigma^2)$ - fordelte.

For at modellen (1) skal være entydig, må vi ha en restriksjon for α_i - ene. (Uten en slik restriksjon kan vi legge en konstant c til μ og trekke den samme konstanten fra alle α_i - ene uten å endre summen $\mu + \alpha_i$.) I variansanalysen er det vanlig å benytte restriksjonen

$$\sum_{i=1}^{I} \alpha_i = 0 \tag{2}$$

jf. avsnitt 11.3 i boka til Devore & Berk. I punktene a - g antar vi at α_i - ene tilfredsstiller restriksjonen (2).

Vi innfører gjennomsnittene

$$\overline{X}_{i\cdot} = \frac{1}{J}\sum_{j=1}^{J} X_{ij} \tag{3}$$

$$\overline{X}_{\cdot\cdot} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J} X_{ij} \tag{4}$$

og lar μ_{ij} angi forventningsverdien til X_{ij}.

a) Vis at

$$\sum_{i=1}^{I}\sum_{j=1}^{J} \left(X_{ij} - \mu_{ij}\right)^2$$

$$= \sum_{i=1}^{I}\sum_{j=1}^{J} \left(X_{ij} - \overline{X}_{i\cdot}\right)^2 + J\sum_{i=1}^{I} \left(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot} - \alpha_i\right)^2 + IJ\,\left(\overline{X}_{\cdot\cdot} - \mu\right)^2$$

b) Forklar at minste kvadraters estimatorer for μ og α_i - ene er gitt ved

$$\begin{aligned}\widehat{\mu} &= \overline{X}_{\cdot\cdot} \\ \widehat{\alpha}_i &= \overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot}\end{aligned} \tag{5}$$

Merk at $\widehat{\alpha}_i$ - ene gitt ved (5) tilfredsstiller restriksjonen (2).

c) Vis at $\widehat{\mu} \sim N(\mu, \sigma^2/IJ)$.

d) Vis at $\widehat{\alpha}_i \sim N(\alpha_i, k\sigma^2)$ for en passende konstant k. Hva blir k?

Vi ser nå på studien av isoflavoner og beintetthet fra oppgave 1.

e) Beregn gjennomsnittene (3) og (4) for dataene i oppgave 1 og bruk dem til å finne minste kvadraters estimater for μ og α_i - ene.

f) Av variansanalysetabellen fra punkt b i oppgave 1 får du et estimat for σ^2. Hva blir estimatet? Bruk estimatet av σ^2 og resultatene i punktene c og d til å bestemme estimerte standardfeil for $\widehat{\mu}$ og $\widehat{\alpha}_i$.

g) Bruk R til å bestemme $\widehat{\mu}$ og $\widehat{\alpha}_i$ med estimerte standardfeil. Kontroller at du får samme resultat som i punktene e og f. (Notatet "Variansanalyse og lineær regresjon" gir hjelp i bruk av R. Notatet er lagt ut på kurssiden.)

Et alternativ til (2) er restriksjonen

$$\alpha_1 = 0 \tag{6}$$

h) Bruk R til å estimere μ og α_i - ene nå vi bruker restriksjonen (6). Forklar hvordan estimatene kan uttrykkes ved gjennomsnittene (3).

LYKKE TIL!