Ekstraoppgaver for STK2120

Geir Storvik

Vår 2011

Ekstraoppgave 1
Anta X_1 og X_2 er uavhengige med $X_1 \sim N(1.0, 1.0)$ og $X_2 \sim N(2.0, 1.5)$. La $\boldsymbol{X} = (X_1, X_2)^T$. Definer

$$\boldsymbol{c} = \begin{pmatrix} 2.0 \\ 3.0 \end{pmatrix}, \quad \boldsymbol{A} = \begin{pmatrix} 1.0 & 0.5 \\ 0.0 & 1.0 \end{pmatrix}$$

For de følgende oppgaver, bruk definisjone på forventning til en vektor og kovariansmatrisen til en vektor til å løse oppgavene.

(*a*) Finn forventningsvektoren til $\boldsymbol{Z} = \boldsymbol{c} + \boldsymbol{A}\boldsymbol{X}$.

(*b*) Finn kovariansmatrisen til $\boldsymbol{Z} = \boldsymbol{c} + \boldsymbol{A}\boldsymbol{X}$.

(*c*) Finn også forventningen til $\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}$.
Hint: Regn ut $\boldsymbol{X}^T\boldsymbol{A}\boldsymbol{X}$ og bruk så egenskaper ved X_1 og X_2.

Ekstraoppgave 2
En tilfeldig vektor $\boldsymbol{X}$ av lengde p er *multivariat normal* fordelt hvis og bare hvis $\boldsymbol{a}^T\boldsymbol{X}$ er univariat normalfordelt for enhver p-vektor a. Vi sier $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ hvis i tillegg $\boldsymbol{X}$ har forventningsvektor $\boldsymbol{\mu}$ og kovariansmatrise $\boldsymbol{\Sigma}$.

(*a*) Anta $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Vis at $X_i \sim N(\mu_i, \Sigma_{ii})$.

(*b*) Anta $\boldsymbol{U}$ er en vektor av uavhengige standard normalfordelte variable. La $\boldsymbol{A}$ være en $p \times p$ matrise slik at $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T$. Vis at

$$\boldsymbol{X} = \boldsymbol{\mu} + \boldsymbol{A}\boldsymbol{U}$$

er multivariat normalfordelt med forventningsvektor $\boldsymbol{\mu}$ og kovariansmatrise $\boldsymbol{\Sigma}$.

Generelt vil det være mange mulige valg av $\boldsymbol{A}$ slik at $\boldsymbol{\Sigma} = \boldsymbol{A}\boldsymbol{A}^T$. Siden $\boldsymbol{\Sigma}$ er symmetrisk og positiv definitt, er en mulighet *Cholesky* dekomposisjonen, som begrenser $\boldsymbol{A}$ til å være nedre triangulær. I **R** kan Cholesky dekomposisjonen av $\boldsymbol{\Sigma}$ finnes ved kommandoen `A=t(chol(Sigma))` (den ekstra `t()` operatoren er nødvendig da **R** beregner $\boldsymbol{A}^T$).

(*c*) I **R** kan uavhengige standard normalfordelte variable enkelt genereres ved kommandoen `rnorm`. Bruk (*b*) til å lage en **R** rutine som genererer en multivariat normfordelt variabel.

Hint: Matrise multiplikasjon gjøres ved operatoren `%*%`.

(*d*) La nå $p = 3$ og

$$\boldsymbol{\mu} = \begin{pmatrix} 1.0 \\ 1.3 \\ 2.0 \end{pmatrix}, \quad \boldsymbol{\Sigma} = \begin{pmatrix} 1.0 & 0.5 & 0.5 \\ 0.5 & 1.0 & 0.5 \\ 0.5 & 0.5 & 1.0 \end{pmatrix}.$$

Simuler 1000 vektorer fra $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ fordelingen og lagre simuleringene i en 1000×3 matrise `m` med hver rad korresponderende til én simulering fra $N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ fordelingen.

Forventningsvektor og kovariansmatrise kan estimeres ved kommandoene

```
colMeans(m)
var(m)
```

Sammenlign disse estimatene med de sanne parameterverdier.

Ekstraoppgave 3 (Genetic linkage modeller)
Anta $n = 197$ dyr (Y) er fordelt i fire grupper som følger:

$$\boldsymbol{x} = (x_1, x_2, x_3, x_4) = (125, 18, 20, 34).$$

Vi vil anta at $\boldsymbol{y}$ følger en *multinomisk fordeling*, dvs

$$p(x_1, x_2, x_3, x_4) = \frac{n!}{(x_1!)(x_2!)(x_3!)(x_4!)} p_1^{x_1} p_2^{x_2} p_3^{x_3} p_4^{x_4}$$

der

$$\boldsymbol{p} = (p_1, p_2, p_3, p_4) = \left(\tfrac{1}{2} + \tfrac{\theta}{4}, \tfrac{1}{4}(1-\theta), \tfrac{1}{4}(1-\theta), \tfrac{\theta}{4}\right) = \tfrac{1}{4}(2+\theta, 1-\theta, 1-\theta, \theta)$$

Vår interesse vil være i å estimere θ.

(*a*) Bestem hvilke verdier θ kan anta.

(*b*) Lag en funksjon i **R** som beregner log-likelihooden for en gitt verdi av θ.

Bruk denne funksjonen til å plotte log-likelihood funksjonen.

(*c*) Beregn skår-funksjonen $s(\theta) = l'(\theta)$ og vis at likningssystemet $s(\theta) = 0$ er ekvivalent med en 2. grads likning.

Finn de to løsningene til denne likningen og argumenter for at kun en av disse løsningene er relevante.

(*d*) Beregn den observerte informasjonen $J_n(\theta) = -l''(\theta)$.

Beregn også Fisher-informasjonen $I_n(\theta) = E[J_n(\theta)]$ og bruk dette til å angi en tilnærmet standardfeil for $\hat{\theta}$.

Argumenter hvorfor $1/\sqrt{J_n(\hat{\theta})}$ er et alternativt estimat på standardfeilen for $\hat{\theta}$. Beregn også denne og sammenlikn.

(*e*) Lag et tilnærmet 95% konfidensintervall for θ.

Vi vil i det etterfølgende se på metoder for numerisk optimering av (log-)likelihood funksjonen. Gitt at vi kan finne maksimum likelihood estimatet mer direkte i dette tilfellet, er det ikke behov for numeriske metoder. Vi vil imidlertid bruke eksemplet som en illustrasjon og gjøre numerisk optimering likevel.

(*f*) Lag en rutine som utfører Newton-Raphson algoritmen. La funksjonen ha input variable initialverdi på θ og observasjonene x. Kjør algoritmen ulike startverdier av θ og se hvordan algoritmen oppfører seg.

(*g*) Lag også en rutine som utfører Fisher-skåring algoritmen.

Kjør også denne rutinen for ulike startverdier.

Hvordan fungerer denne algoritmen i forhold til Newton-Raphson i dette eksemplet?

Ekstraoppgave 4 (Egenskaper ved kovariansmatriser)
En $p \times p$ symmetrisk matrise $\boldsymbol{M}$ er positiv semidefinitt hvis $\boldsymbol{a}^T \boldsymbol{M} \boldsymbol{a} \geq 0$ for enhver p-vektor $\boldsymbol{a}$. Matrisen er positiv definitt hvis $\boldsymbol{a}^T \boldsymbol{M} \boldsymbol{a} > 0$

(a) Vis at en diagonal matrise er positiv semidefinitt hvis og bare hvis diagonal-elementene er ikke-negative.

(b) Vis at $\boldsymbol{M}$ er positiv semidefinitt hvis og bare hvis alle egenverdiene til $\boldsymbol{M}$ er ikke-negative.

Hint: En symmetrisk matrise $\boldsymbol{M}$ kan alltid skrives på formen $\boldsymbol{M} = \boldsymbol{\Gamma}\boldsymbol{\Lambda}\boldsymbol{\Gamma}'$ der $\boldsymbol{\Lambda}$ er en diagonal matrise av egenverdiene til $\boldsymbol{M}$ og $\boldsymbol{\Gamma}$ er en ortogonal matrise med standardiserte egenvektorer som kollonner.

(c) Hva blir det tilsvarende resultatet for positiv definitte matriser?

Anta nå $\boldsymbol{X}$ er en tilfeldig vektor med forventningsvektor $\boldsymbol{\mu}_x$ og kovariansmatrise $\boldsymbol{\Sigma}_x$.

(d) Vis at $\boldsymbol{\Sigma}_x$ alltid er symmetrisk.

(e) Vis at $\boldsymbol{\Sigma}_x$ alltid er positiv semidefinitt.
Hint: Bruk at variansen til $\boldsymbol{a}^T \boldsymbol{X}$ må være ikke-negativ.

(f) I hvilke tilfeller vil $\boldsymbol{\Sigma}_x$ ikke være positiv definitt?

(g) Hvorfor er Fisher's informasjonsmatrise alltid positiv semidefinitt?

Ekstraoppgave 5 (Poisson regresjon)
For telledata blir vanligvis Poisson fordelingen gitt ved

$$f(y; \lambda) = \frac{\lambda^y e^{-\lambda}}{y!}$$

benyttet. Vi skal i denne oppgaven se på situasjoner der intensitetsparameteren λ varierer for de ulike observasjoner men der variasjonene (forhåpentligvis) kan forklares ved noen forklaringsvariable $\boldsymbol{x}$. Som illustrasjon vil vi bruke et datasett som angir antall barn for hver av 141 gravide kvinner samt alder på kvinnene. Disse data kan leses fra kursets hjemmeside.

Vi vil imidlertid starte med å se på den generelle modellen

$$Y_i \sim \text{Poisson}(\lambda(\boldsymbol{x}_i))$$

der

$$\lambda(\boldsymbol{x}_i) = \exp(\sum_{j=0}^{p} x_{i,j}\beta_j)$$

og der $x_{i,0} = 1$. Denne modellen kalles *Poisson regresjonsmodellen.*

(*a*) Vis at log-likelihood funksjonen $l(\boldsymbol{\beta})$ for observasjoner $\{(\boldsymbol{x}_i, y_i), i = 1, ..., n\}$ er gitt ved

$$l(\boldsymbol{\beta}) = \sum_{i=1}^{n}[y_i(\sum_{j=0}^{p} x_{ij}\beta_j) - \exp(\sum_{j=0}^{p} x_{ij}\beta_j) - \log(y_i!)]$$

(*b*) Beregn skår funksjonen og vis at hvert element i denne har forventning lik 0.

(*c*) Vis at den observerte informasjonsmatrisen $\bar{\boldsymbol{J}}(\boldsymbol{\beta})$ er gitt ved

$$\bar{\boldsymbol{J}}_{k,l} = \sum_{i=1}^{n} \exp(\sum_{j=0}^{p} x_{ij}\beta_j)x_{ik}x_{il}$$

(*d*) Finn Fisher informasjonsmatrisen $\bar{\boldsymbol{I}}(\boldsymbol{\beta})$.
Vis også at $\bar{\boldsymbol{I}}(\boldsymbol{\beta})$ alltid er positiv definitt.

(*e*) Implementer en rutine som beregner maksimum likelihood estimatene basert på Fisher-skåring algoritmen med en generell vektor $\boldsymbol{x}_i$ av forklaringsvariable.

Hint: Det kan være lurt å legge inn en kollonne med 1-ere i matrisen $\boldsymbol{x}$ ved `x = cbind(1,x)` helt i starten av rutinen.

(*f*) Se nå på datasettet med antall barn. Tilpass en modell der

$$\lambda(x_i) = \exp(\beta_0 + x_i\beta_1).$$

Finn ML estmater og konstruer tilnærmede 95% konfidensintervaller for β_0 og β_1. Kjør rutinen på dataene beskrevet ovenfor. Finn estimater og lag 95% konfidensintervaller for β_0 og β_1.

(*g*) Utvid så modellen til

$$\lambda(x_i) = \exp(\beta_0 + x_i\beta_1 + x_i^2\beta_2).$$

Finn også ML estimater for denne modellen. Basert på et 95% konfidensintervall for β_2, synes du det er et poeng i å ha med et kvadratisk ledd her?

Ekstraoppgave 6 (Estimater av kummulative fordelinger)
La $f(x)$ være en sannsynlighetstetthet over den reelle akse og la

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)du.$$

være den kummulative fordelingsfunksjon. Innen bootstrapping er et estimat på $F(x)$ essentielt. Gitt et tilfeldig utvalg $\boldsymbol{x} = (x_1, ..., x_n)$ fra F, så er et mye brukt estimat i denne forbindelse den *empirisk kummulative fordeling* definert ved

$$\widehat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \leq x) \qquad (*)$$

der $I(A)$ er indikatorfunksjonen, dvs $I(A) = 1$ hvis begivenheten A intreffer og 0 ellers. Vi skal i denne oppgaven studere egenskaper til $\widehat{F}(x)$.

(a) Følgende kommandoer simulerer $n = 15$ verdier fra normalfordelingen med forventning $\mu = 1$ og standard avvik $\sigma = 2$ og plotter $\widehat{F}$ sammen med F:

```
x <- rnorm(15,2,1)
plot(ecdf(x),verticals=TRUE)
lines(sort(x),pnorm(sort(x),2,1),lty=2)
```

Prøv ut disse kommandoene gjentatte ganger og se på hvordan $\widehat{F}$ varierer.

(b) Vis at $E[I(A)] = P(A)$ der $P(A)$ er sannsynligheten for begivenhet A.

(c) Vis at $E[\widehat{F}(x)] = F(x)$ for alle verdier av x.

(d) Vis at $\text{var}[\widehat{F}(x)] = \frac{1}{n}F(x)[1 - F(x)]$.

Hint: Vis først at $\text{var}[\widehat{F}(x)] = \frac{1}{n}\text{var}[I(X \leq x)]$ der $X \sim F$.

(e) Betrakt nå den diskrete sannsynlighetsfordelingen gitt ved $p(x_i) = P(X = x_i) = \frac{1}{n}$ for $i = 1, ..., n$. Vis at den kummulative fordelingsfunksjonen til denne fordelingen er lik $\widehat{F}$.

La oss nå se på en situasjon der vi antar $f(x)$ tilhører en parametrisk familie beskrevet ved $f(x;\theta)$ med tilhørende kummulativ fordeling

$$F(x;\theta) = P(X \leq x;\theta) = \int_{-\infty}^{x} f(u;\theta)du.$$

Et alternativt estimat i denne situasjonen er

$$\widehat{F}(x) = F(x;\hat{\theta})$$

der $\hat{\theta}$ er et estimat for θ (f.eks. maksimum likelihood estimatet).

(f) Følgende kommandoer simulerer $n = 15$ verdier fra normalfordelingen med forventning $\mu = 1$ og standard avvik $\sigma = 2$ og plotter det alternative estimatet $\widehat{F}$ sammen med F:

```
x <- rnorm(15,2,1)
mu.hat = mean(x);sigma.hat = sqrt(var(x))
plot(sort(x),pnorm(sort(x),mu.hat,sigma.hat),type="l")
lines(sort(x),pnorm(sort(x),2,1),lty=2)
```

Prøv ut disse kommandoene gjentatte ganger og se på hvordan $\widehat{F}$ varierer.

Diskutér fordeler/ulemper med dette parametriske estimatet av F kontra den ikke-parametriske versjonen (*).

Ekstraoppgave 7 (Bootstrap og gjennomsnitt)
La F være en kummulativ fordelingsfunksjon og f dens deriverte (dvs sannsynlighetstettheten). Anta $\theta(F) = \int_x f(x)dx$ der vi har spesifisert at θ avhenger av F gjennom notasjonen $\theta(F)$. Vi vil i det etterfølgende også bruke notasjonen $\theta(F) = \mathsf{E}^F[X]$ der E^F betyr forventningen mhp fordelingen spesifisert av F. La videre $\hat{\theta} = \hat{\theta}(\boldsymbol{x}) = \bar{x}$ være en estimator for $\theta(F)$ basert på et tilfeldig utvalg $\boldsymbol{x} = (x_1, ..., x_n)$ fra fordelingen F.

(a) Vis at hvis

$$\widehat{F}(x) = \frac{1}{n}\sum_{i=1}^{n} I(x_i \leq x)$$

så er

$$\mathsf{E}^{\widehat{F}}[\hat{\theta}] = \mathsf{E}^{\widehat{F}}[X] = \bar{x}$$

dvs forventningen i fordelingen $\widehat{F}$ er $\bar{x}$.

Hint: Bruk resultatet fra ekstraoppgave 6(e).

(b) Vis at $\theta(\widehat{F}) = \bar{x}$ samt at bootstrap skjevheten for $\hat{\theta}$ er 0.

(c) Vis at variansen i fordelingen $\widehat{F}$ er

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

og bruk dette til å finne variansen til $\hat{\theta}(\boldsymbol{x}^*)$ der $\boldsymbol{x}^* = (x_1^*, ..., x_n^*)$ er et tilfeldig utvalg fra fordelingen $\widehat{F}$. Hva blir dermed bootstrap estimatet for variansen til $\hat{\theta}$?

Ekstraoppgave 8 (Efron & Tibsherani, oppgave 6.10)
Betrakt det kunstige datasettet bestående av 8 tall

$$1.2, 3.5, 4.7, 7.3, 8.6, 12.4, 13.8, 18.1$$

La $\hat{\theta}$ være det 25% trimmede gjennomsnitt, beregnet ved å slette de to laveste og de to høyeste tallene og så ta gjennomsnittet av de resterende fire tall

(*a*) Estimer ved hjelp av boostrapping standardfeilen til $\hat{\theta}$ for $B = 25, 100, 200, 500, 1000, 2000$. Fra disse resultatene, estimater det "ideelle" bootstrap estimat $s_{\hat{\theta}}$ svarende til $B = \infty$.

(*b*) Repeter (*a*) 10 ganger for å anslå usikkerheten i estimatene på standardfeilen. Hvor stor må B være for å få en akseptabel nøyaktighet?

Ekstraoppgave 9 (Nedbør i Illinois)
Vi vil i denne oppgaven igjen betrakte observasjonene over nedbørsmengde i Illinois.

(*a*) I Bootstrap notatet ble forventningsskjevet og standard feil til estimatene $\hat{\lambda}$ og $\hat{\alpha}$ basert på parametrisk bootstrapping diskutert. Lag nå bootstrap konfidensintervaller for α og λ i gamma fordelingen basert på nedbørsdataene.

(*b*) La oss nå løsrive oss fra antagelsen om Gamma fordeling. Parametrene α, λ er da ikke så meningsfulle lenger. Vi vil derfor istedet se på *spredningskoeffisienten*, definert ved

$$C = \frac{\sigma}{\mu}$$

Et naturlig estimat for denne er

$$\hat{C} = \frac{s}{\bar{x}}$$

der s er det empiriske standard avviket. Bruk bootstrapping til å si noe om forventningsskjevhet og standard feil til $\hat{C}$ og til å konstruere et 95% konfidensintervall for C.

Ekstraoppgave 10 (Bootstrapping og regresjon)
Vi vil i denne oppgaven se på hvordan bootstrapping kan utføres i forbindelse med regresjon. La $\hat{\beta}_0$ og $\hat{\beta}_1$ være minste-kvadraters estimater fra en lineær regresjonsmodell

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

Som illustrasjon vil vi bruke data fra oppgave 12.74 i Devore-Berk boka.

(*a*) Utfør først en vanlig lineær regresjon ved bruk av `lm` rutinen i **R** . Sjekk residualene om normalitet er rimelig.

Vi ønsker nå å se på forventningsskjevhet, standardfeil og konfidensintervaller når vi ikke lenger antar at ε_i-ene er normalfordelte.

(b) Anta først nye $(x_i^*, y_i^*), i = 1, ..., n$ blir trukket på følgende måte:

(i) Sett $x_i^* = x_i, i = 1, ..., n$.

(ii) Repeter for $i = 1, ..., n$: Trekk tilfelding ε_i^* med tilbakelegging fra $\{\hat{\varepsilon}_1, ..., \hat{\varepsilon}_1\}$ der $\hat{\varepsilon}_i = y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i$.

(iii) Sett $y_i = \hat{\beta}_0 + \hat{\beta}_1 x_i + \varepsilon_i^*$ for $i = 1, ..., n$.

Bruk denne prosedyren til å utføre bootstrapping på de gitte data. Sammenlikn resultatene med de du fikk da normalfordelingsantagelsen ble brukt (dvs resultatene fra (a)).

(c) En alternativ måte å trekke $(x_i^*, y_i^*), i = 1, ..., n$ på er følgende:

(i) Trekk $j_1^*, ..., j_n^*$ tilfeldig med tilbakelegging fra $\{1, ..., n\}$.

(ii) Sett $(x_i^*, y_i^*) = (x_{j_i}, y_{j_i})$ for $i = 1, ..., n$.

Prøv igjen på de samme data. Diskuter forskjeller/likheter med resultatene i (a).

De to alternative prosedyrene relaterer seg til om vi oppfatter forklaringsvariablene (x_i-ene) som tilfeldige eller ikke. I planlagte forsøk, der vi selv kan bestemme x_i-ene, er det mest naturlig å tenke på forklaringsvariablene som faste. I mange observasjonsstudier samles både x_i og y_i inn og det er mer naturlig å se på x_i-ene som tilfeldige.

(d) Hvilke av de to prosedyrer er mest naturlige å bruke i når x_i-ene betraktes som faste? Hva hvis de er tilfeldige?

(e) Hvis vi er usikre på om den lineære strukturen $E[y_i] = \beta_0 + \beta_1 x_i$ er riktig, hvilken av prosedyrene er mest robust i forhold til denne antagelsen?

Ekstraoppgave 11 (Testing i multinomiske fordelinger)
Anta $(N_1, ..., N_k)$ er multinomisk fordelt med n forsøk og sannsynlighet p_i for kategori i,

$$p(N_1 = n_1, ..., N_k = n_k) = \frac{n!}{n_1! \cdots n_k!} \prod_{i=1}^{k} p_i^{n_i}.$$

Vi ønsker å teste

$$H_0 : p_i = p_{i0}, i = 1, ..., k \text{ mot } H_a : p_i \neq p_{i0} \text{ for minst en } i.$$

I kapittel 13 i læreboka beskrives χ^2-observatoren som er gitt ved

$$\chi^2 = \sum_{i=1}^{k} \frac{(n_i - np_{i0})^2}{np_{i0}} = \sum_{i=1}^{k} \frac{(n_i - E_i)^2}{E_i}$$

der E_i er brukt for forventet antall i kategori i.

Gitt det vi har lært i kapittel 9, avsnitt 9.5, ville det kanskje være mer naturlig å bruke LR-observatoren (LR=likelihood ratio). I denne oppgaven vil vi vise at i praksis har det ikke så mye betydning hvilken vi bruker (bare vi har en viss mengde data). At det er blitt mest vanlig å bruke χ^2 har nok mest historiske årsaker.

La $\theta = (p_1, ..., p_k)$ og $\Omega = \{(p_1, ..., p_k); 0 \leq p_i \leq 1, \sum_i p_i = 1\}$.

(a) Vis at ML estimatene innenfor Ω er gitt ved $\hat{p}_i = n_i/n$.

(b) Finn LR-observatoren og vis at

$$-2\log(LR) = 2\sum_{i=1}^{k} n_i \log\left(\frac{n_i}{E_i}\right) = \sum_{i=1}^{k} f_{E_i}(n_i)$$

der

$$f_{x_0}(x) = x \log\left(\frac{x}{x_0}\right).$$

(c) Vis at for x i nærheten av x_0 så er

$$f_{x_0}(x) \approx (x - x_0) + \frac{1}{2x_0}(x - x_0)^2.$$

(d) Vis at for n stor så er

$$-2\log(LR) \approx 2n \sum_{i=1}^{k} [\hat{p}_i - p_{i0}] + n \sum_{i=1}^{k} \frac{(\hat{p}_i - p_{i0})^2}{p_{i0}} = \chi^2.$$

(e) Vis så at de tilnærmede fordelinger for $-2\log(LR)$ og χ^2 under H_0 er like.

(f) For data i oppgave 13.3, beregn begge testobservatorene samt tilhørende P-verdier og sammenlikn.