|  |
| --- |
| Grizli777 |
| Homework 1 |
| Statistische modellen en data analyse |
|  |
| **Andreas Put – Ma1 Computerwetenschappen** |
| **27/04/2011** |

|  |
| --- |
| [Type the abstract of the document here. The abstract is typically a short summary of the contents of the document. Type the abstract of the document here. The abstract is typically a short summary of the contents of the document.] |

# Vraag 1

## deel 1

FX(x) =

De dichtheidsfunctie fX(x) van FX(x) kan berekend worden door Fx af te leiden naar x

fX(x) =

= - ln(1 − p)

Via de formule MLE = kunnen we de maximum likelyhood estimator vinden

MLEp = argmax

# Vraag 2

## Deel 1

De onderstaande formules worden veelvuldig gebruikt voor deze vraag:

E[X2|X1=x1] = µ2 + Σ21 Σ11-1 (x1 - µ1) (1) µ =

Cov[X2|X1 = x1] = Σ22 - Σ21 Σ21-1 Σ12  (2)Σ=

We beginnen bij Y|Z  
Y|Z is normaal verdeeld: Y|Z ~ N( 1+Z , 1) (3)  
Vullen we x1 = µ1 in in (1) dan bekomen we  
E[X2|X1= µ1] = µ2 = 1+Z (3). Omdat Z = x1 = µ1 = 0 🡺 µ2 = 1  
Vullen we x1 = 2 in en daar Σ11 = 1 bekomen we  
E[X2|X1=2] = 1+Z = 3 = 1 + Σ21(2-0) 🡺 Σ21 = 1 = Σ12

Σ22 verkrijgen we via (2)  
Cov[X2|X1 = 0] = 1 = Σ22 – 1/1\*1 🡺 Σ22 = 2

* [Z Y]T ~ ( , )

Vervolgens analyseren we X|Y,Z  
Vullen we x1 = µ1 = [0 1]T in in (1) dan bekomen we E[X2|X1= µ1] = µ2 = 1-Y = 0  
Vullen we x1 = [1,1]T in en daar Σ11 = is, bekomen we  
E[X2|X1= x1] = 1-Y = 0 = Σ21 = Σ21 = 0 (4)  
Met x1 = [0 0] T bekomen we E[X2|X1= x1] = 1-Y = 1 = Σ21 = Σ21 = 1 (5)

Uit (4) en (5) kunnen we Σ21 halen 🡺 Σ21 = [1 2] = Σ12T

Σ22 verkrijgen we via (2)  
Cov[X2|X1 = µ1] = 1 = Σ22 – Σ12 Σ11-1 Σ12T🡺 Σ22 = 3

* [Z Y X]T ~ ( , ) 🡺 [X Y Z]T ~ ( , )

## Y|X,Z

Hiervoor is de verdeling van [X,Z]T nodig:  
[X Z]T ~ ( , ) X,Z,Y ~ ( , )

E[Y|X,Z = µ] = 1 + [2 1] -1) = ½(Z+X)  
Cov[Y|X,Z = µ] = 2- [2 1] -1) = 2 - 1.5 = 0.5

Y|X,Z ~ (½(Z+X) , 0.5)

## Deel 2

De verdeling van [U V ]T waar U = 1 + Z en V = 1 - Y :  
U = 1 + Z 🡺 µu = 1 + µz ; Su = Sz 🡺 U ~ (1,1) - Er verandert niets aan de variantie na optelling van een constante.

V = 1 – Y 🡺 µv = 1 - µy ; Sv = Sy  🡺 V ~ (0,2)

µU,V = [ 1 0 ]

ΣU,V = =

## Deel 3

Omdat de variantie van U gelijk is aan die van Z, en de variantie van V is gelijk aan de variantie van Y, is de covariantiematrix van [U,Y] gelijk aan die van [U,V]

E(Y|U = 2) = µY + Σ21 Σ11-1 (2 - µU) ~( , )  
 = 1 + 1/1 \* (2-1) = 2

# Vraag 3

Gegeven de steekproef met verdeling

*N3* ~

De hypothese H0 : µ1 = µ2 = µ3 v.s. H1 : µi≠µj kan als volgt geëvalueerd worden.  
We stellen een alternatieve hypothese op:

H0' : µ = = 0 v.s. H1' : µ ≠ 0

Via de T2-test van Hotelling kan deze hypothese aangenomen of verworpen worden.

Omdat deze test affien invariant is, zal de volgende hypothese

HY0 : Aµ = 0 v.s. HY1 : Aµ ≠ 0 (det(A)≠0)

hetzelfde resultaat opleveren. De kritieke waarde van de T2 test zal immers hetzelfde zijn als bij de originele hypothese test.  
Een geschikte matrix A hiervoor is: A1 = (det(A1) = 1)  
A1\*µ levert als resultaat:

Een andere matrix die we kunnen opstellen is: A2 = det(A2) = 1  
A2\*µ levert als resultaat:   
Door het toepassen van de transformaties door de matrix Ai wordt getest of µ1 = µ2 = µ3.

# Vraag 4

## Deel 1

De R code voor deze oefening is te vinden in bijlage ....

Omdat we (ongecorreleerde) gepaarde steekproeven behandelen, is het aantal genomen samples gelijk aan n = 10

De hypothesetest

H0 : µ1 = µ2 v.s. H1 : µ1 ≠ µ2

kan worden opgelost door eerst de twee gepaarde steekproeven samen te voegen. De kenmerken van de resulterende steekproef worden dan:

Xp = X1 – X2 en Sp  =

Met deze samengestelde gemiddelde en covariantiematrix kan een nieuwe hypothese worden opgesteld:

H’0 : µp = 0 v.s. H1 : µ1 ≠ 0

De test of de hypothese geldt of niet wordt uitgevoerd door middel van de Hotelling test voor gepaarde steekproeven.

T2 = n \* xpT \* Sp-1 \* xp

En zo kan de F waarde berekend worden:

F =

Als deze waarde groter is dan de critische waarde F0.95,2,8 zal de hypothese verworpen worden.

In dit geval geldt : F =7.901235, F0.95,2,8= 4.45897  
(Alle tussenberekeningen en uitkomsten kunnen verkregen worden door het R bestand uit te voeren)

De hypotheektest leert ons dat µp ≠ 0, en dus µ1 ≠ µ2, maar het zegt niets over welke elementen hiervoor verantwoordelijk zijn.

## Deel 2

Om dit te onderzoeken gebruiken we betrouwbaarheidsintervallen voor de verschillende elementen van xp. Wanneer 0 niet in dit betrouwbaarheidsinterval ligt, zal de nulhypothese verworpen worden. Omdat Xp = X1 – X2 zal als 0 niet in het betrouwbaarheidsinterval van Xpi ligt ook X1i ≠ X2i zijn.

Het betrouwbaarheidsinterval voor Xp(i) wordt via de volgende formule berekend:

[Xp(i) ± tn-1,1-α/2]

Voor H0 : µ11 = µ21 v.s. H1 : µ11 ≠ µ21 bekomen we het volgende betrouwbaarheidsinterval [0.2477406 3.7522594]

Omdat 0 niet in dit interval ligt, kunnen we de nulhypothese verwerpen.

Voor H0 : µ12 = µ22 v.s. H1 : µ12 ≠ µ22 bekomen we het volgende betrouwbaarheidsinterval

[-1.752259 1.752259]

0 ligt in dit interval, en we nemen de hypothese dus aan.

######EXERCISE 4######

###Part 1###

matrix(c(3,1),ncol = 1)->x1

matrix(c(1,1),ncol = 1)->x2

matrix(c(4,-1 ,-1,2),nrow = 2, ncol=2, byrow=TRUE)->s1

matrix(c(2,-2 ,-2,4),nrow = 2, ncol=2, byrow=TRUE)->s2

#Merge the covariance matrices and means

sp <- (9\*s1+9\*s2)/(18);sp

xp <- x1-x2;xp

#T² Hotelling test

#calculate T²

t2<- 10\*t(xp)%\*%(solve(sp))%\*%(xp);t2

#Calculate F value

f<-(10-2)\*t2/((10-1)\*2);f #f <- (n1+n2-k-1)\*t2/((n1+n2-2)\*k)

#Calculate critical value

crit<-qf(.95,2,8)

#regection regian

c(crit,Inf) ## if f in rejection region ==> Reject!

###Part 2###

#Confidence interval for x(1,1)

#If 0 lies in the CI, we accept the hypothesis, otherwise reject.

t <- qt(.975,9);t

cil1 <- xp[1] - t\*sqrt((2\*sp[1,1])/10)

cih1 <- xp[1] + t\*sqrt((2\*sp[1,1])/10)

c(cil1,cih1)

###Part 3###

#Confidence interval for x(2,1)

#If 0 lies in the CI, we accept the hypothesis, otherwise reject.

cil2 <- xp[2] - t\*sqrt((2\*sp[2,2])/10)

cih2 <- xp[2] + t\*sqrt((2\*sp[2,2])/10)

c(cil2,cih2)