STAT 802 Multivariate Analysis Sept. 18, 2003

Today: More details from Ch 4:

Bivariate Case p 151

Distribution of Lin Combination of univariate (possibly correlated) normals p 157

Chi-Square Distribution of Quadratic Forms based on $\Sigma^{-1}$ p 163-4

Bivariate Normal ($\rho \neq 0$ say)

How do you simulate it?

One way: generate 3 IID N(0,1) variates $z_1$ $z_2$ $z_3$

Let $x=z_1+a\ z_3$ and $y=z_2 + a\ z_3$ where $a=(\rho/(1-\rho))^{1/2}$

Program “bnorm” at end of notes:

Look at data scatter – Does it look like multivariate normal?

Can you imagine the density contours?

A program to look at the scatter and estimated density contours – see program “density2D” at end of notes (Disclaimer – it works, but the code is not optimal!)

Exercise: Devise a general method for simulating a p-variate normal distribution for a given covariance matrix – w.l.o.g., assume means = 0. (One way is to use eigenanalysis! You can test your method using the bivariate program above. If you assume that all the covariances are positive, you can use equation (2.22) for the square root transformation with A replaced by the desired $\Sigma$.)

P 157: Let A be q x p. If $X \sim N_p(\mu,\Sigma)$ then $AX \sim N_q(A\mu,A\Sigma A')$.

If $Z \sim N_p(0,I)$, then AZ is $N_q(0,AA')$. So the task in the above exercise can be solved by finding A so $AA' = \Sigma$.

P 163 If If $X \sim N_p(\mu,\Sigma)$, then $(X-\mu)'\ \Sigma^{-1}\ (X-\mu) \sim$ Chi-square on p df. Moreover, the solid ellipsoid $\{x: (x-\mu)'\ \Sigma^{-1}\ (x-\mu) \leq \chi_p^2(\alpha)\}$ contains $1-\alpha$ of the probability.

It would be interesting to demonstrate this – simulate $X \sim N_p(\mu,\Sigma)$ and for various $\alpha$ compute the proportion in the ellipsoidal set. (How?)

P 182: Correlation criterion on Normal Q-Q plot as a test of normality.

Try with simulated normal data.

P 185-7: A Chi square probability plot for squared distances whwn X is Normal.

But see warning on top of page 189 about futility of testing GOF.

Gmacro

```
bnorm.mac
let k1=.5     #desired correlation
let k2=10     #desired sample size
let k3=(k1/(1-k1))**.5
rand k2 c1-c3
let c1=c1+k3*c3
let c2=c2+k3*c3
corr c1 c2
name k1 'nominalr'
print k1
endmacro
```

---

```
Gmacro
Density2D.mac
#MINITAB  program  to produce density  plots  from
#x=c1 y=c2  c1 is horizontal, c2 is vertical
layout;
aspect 1 1.
plot c2*c1    # just to visualize the bivariate distribution of sample points
endlayout
brief 0      # this  avoids output so speeds up calculation
mini c1 k18  # determine range of data to set smoothing constant
maxi c1 k19
let k20=k19-k18  #range of c1
mini c2 k21
maxi c2 k22
let k23=k22-k21   # range of c2
n c1 k24         # how many points?
let k8=k20/(k24**.5)   # This is the smoothing constant - big is smooth -
for c1=x
let k15=k23/(k24**.5)   # smoothing const for c2=y
let k16=11   #first column for  density estimates
let k4=k16   # this changes during program, wheras k16 stays at intial
value
let k5=10   # number of columns
let k6=10    # number of rows
let k17=k16+k5-1   # this is index of last column generated
mini c1 k9
maxi c1 k11
mini c2 k10
maxi c2 k12
```

```
let k13=(k11-k9)/(k5-1)
let k14=(k12-k10)/(k6-1)
Do k2=1:k5    #first fix a column number
Do k1=1:k6    # go through all the rows of a particular column
let c7=exp(-(((k9+(k2-1)*k13)-c1)/k8)**2-(((k10+(k1-1)*k14)-
c2)/k15)**2)
sum c7 k7     #compute pseudo-density
#let c6=c3*exp(-(((k9+(k2-1)*k13)-c1)/k8)**2-(((k10+(k1-1)*k14)-
c2)/k15)**2)
#sum c6 k3     #compute pseudo z-value
let ck4(k1)=k7  #record  density  in row k1, column  k4
enddo
let k4=k4+1    # go to next column
enddo
stack ck16-ck17 c80   # put density in  one column
set c81
(k9:k11/k13)k6                # create  ID for x grid
end
set c82                   # create ID for y grid
k5(k10:k12/k14)
end
name c80 'density' c81 'c1'  c82 'c2'
layout;            # these commands set up the aspect ratio
aspect 1 1.
ContourPlot c80*c82*c81;
Connect;
Type 1 1 1 1 1 1 1 1;    #type of line
color 4 4 4 4 2 2 2 2;   # color of line
size 3 2.5 2 1 1 2 2.5 3;    #weight of line
nlevel 8.           #number of lines
endlayout          # you need this to get the plot going (after using layout)
brief2
endmacro
```