STAT 400 Data Analysis October 8, 2003

Today: An aside about “error in X” regression
Constructing Smooth Contour Plots

Error-in-X regression: usual estimator tends to bias slope estimate downwards.

Demo with program:

```
Gmacro
errvbls.mac
set c1
25(1),25(2)
end
rand 50 c4 c5
name c1 'X' c2 'Xerr' c3 'Y' c4 'Z1' c5 'Z2' c6 'coeff' c7 'coefferr'
let 'xerr'='x' + 'Z1'
let 'Y'='X'+'Z2'
Regr 'Y' 1 'X';
coeff 'coeff'.
Regr 'Y' 1 'Xerr';
coeff 'coefferr'.
print 'coeff' 'coefferr'
endmacro
```

Explanation:

I argue that the errors away from the mean of X would have a greater influence on the slope than the errors equidistant towards the mean of X.
The effect is to REDUCE the slope of the line from what it would be without errors. The X values without error are associated with certain Y values but if there is error in X these same values of Y will appear to be associated with values of X away from the true values. If the distribution of measurement errors in X is symmetrical, then an observed $X^* = X + |d|$ will occur along with $X-|d|$ to be associated with values of Ycentered at $E(Y|X)$. But the observed Y associated with value of X^* that is farther from the mean of X will tend to reduce the slope from its true value while the value of X^* closer to the mean of X will tend to increase the slope from its true value. However, the effect of the more remote X^* on the slope is greater than the effect of the more central X^*. (To see this, think what happens when the $X^*+-|d|$ is = the mean of the X - no effect at all).
Ergo, the error in variables will cause the usual slope estimation formula to underestimate the slope - bias!

Smooth Contour Plots:

First smooth surface (at grid points), then construct contour lines.

A smoothing program called “surface”.

```
Gmacro
surface.mac
#MINITAB  program  to produce smoothes surface  plots  from
#x=c1 y=c2 z=c3  c1 is horizontal, c2 is vertical, c3 is out from page
brief 0      # this  avoids output so speeds up calculation
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 k8=(k11-k9)/6   # This is the smoothing constant - big is smooth - for
c1=x
let k15=(k12-k10)/6  # smoothing const for c2=y
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)=k3/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
layout;            # these commnands set up the aspect ratio
aspect 1 1.
name c80 'Z' c81 'X' c82 'Y'
ContourPlot c80*c82*c81;
Connect;
Type 1 1 1 1 1 1 1 1 1 1 1 1 1 1;    #type of line
color 4 4 4 4 4 4 4 2 2 2 2 2 2 2;   # color of line
size 4 3.5 3 2.5 2 1.5 1 1 1.5 2 2.5 3 3.5 4;    #weight of line
nlevel 14.            #number of lines
endlayout          # you need this to get the plot going (after using
laoyout)
SurfacePlot c80*c82*c81;
Surface;
EColor 2;
ESize 3;
Irregular;
NMesh 1 15;
NMesh 2 15;
VPosition -5.5 1.5 1.0;
VField 2 4.0;
VAspect 3 3 2;
HSRemoval 2;
LShading 0 1;
Light 2 0 1;
Color 1.
Endmacro
```

But this only gives the surface height at the grid points – how do we use this to draw smooth contour lines? See pp 240-243.