# Solution to Exercise 2: A statistical process control chart

Key points:

- You must determine on how to aggregate data to obtain the numerator and denominator and, where necessary, the time components over the observation period
- For a binomial outcome, a PChart is the appropriate SPC chart

Task

- *Produce five PCharts to display the proportion of 1) positive smears among all smears, 2) low-positive smears among all smears, 3) scanty positive smears among all smears, 4) low-positive smears among all positive smears, and 5) scanty positive smears among all positive smears.*

These are the five PCharts:

Proportion scanty pos smears among all positive smears
Uganda, 1999-2001

Course Part D, Exercise D_EX02, in EpiData Analysis
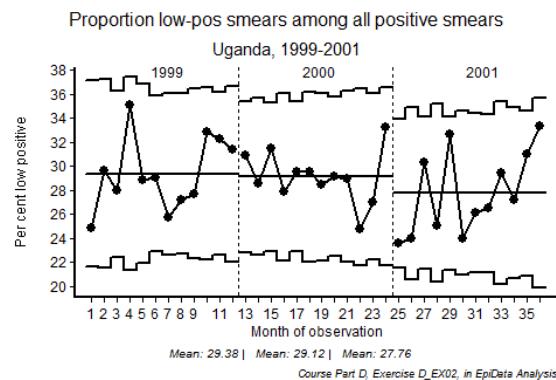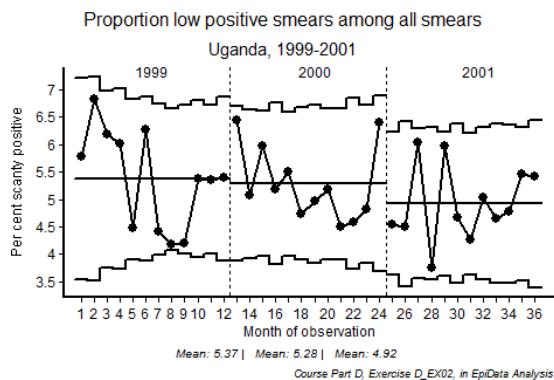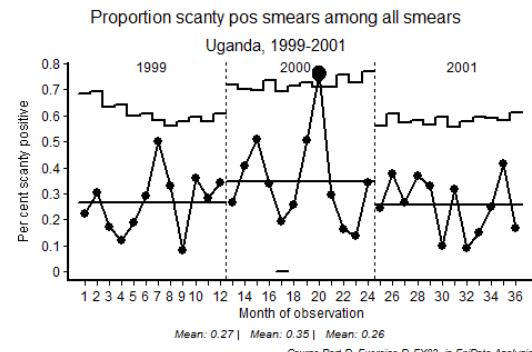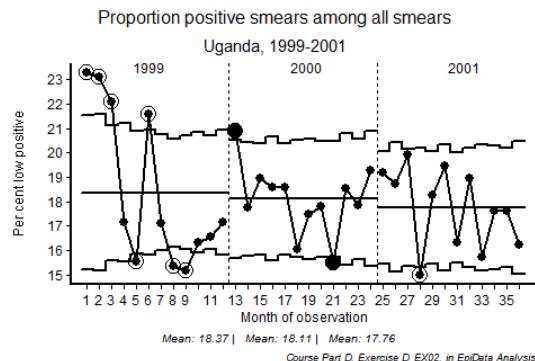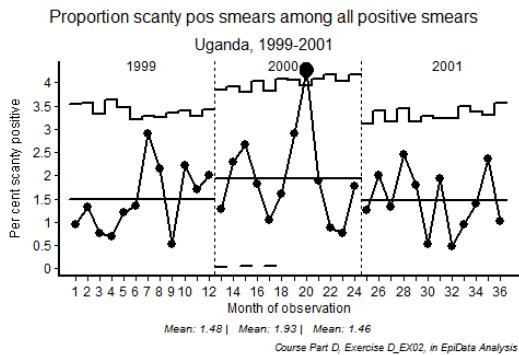
This is the program d_ex02.pgm that produced them:

```
* Part D, Exercise 2

* 1) Determine the proportion of positive
*    smears among all smears
* 2) Determine the proportion of low-postive
*    smears among all positive smears
* 3) Determine the proportion of scanty positive
*    semars among all positive smears
* Definition of positive: any quantified positive
*    or any scanty (quantified scanty or unquantified scanty)
* Definition low-positive: any smear which is 1+ positive or scanty positive

* Written by:    Hans L Rieder
* First version: 26 Jun 2011
* Last revision: 29 Apr 2013

cls
close
logclose

*****************************************************
* Procedural steps
* 1) Make basic dataset
* 2) Start selection process
* 3) Aggregate data
* 4) Make SPC charts

*****************************************************
* 1) Prepare basic dataset

cls
close

read "mmuz.rec"

define regyear ####
regyear=year(regdate)
label regyear "Registration year"

cls
tables country regyear

gen i mmseq=0
if year(regdate)=1999 then mmseq=month(regdate)
if year(regdate)=2000 then mmseq=month(regdate)+12
if year(regdate)=2001 then mmseq=month(regdate)+24
if year(regdate)=2002 then mmseq=month(regdate)+36
if year(regdate)=2003 then mmseq=month(regdate)+48
label mmseq "Sequential month"

select reason=0
select country=3
select mmseq>0 and mmseq<37
* The selection above is not necessary for Uganda alone
* as there are no examinees in 2003

keep result1 result2 result3 regyear mmseq
savedata "temp_01.rec" /replace
```

```
**************************************************
* 2) Count all smears, all positive, all scanty positive,
* all low positive smears

cls
close
read "temp_01.rec"

cls
* Count all smears
* Note: non-sensical sequences removed, thus simply:
gen i allsmears=1
if result2<>9 then allsmears=2
if result3<>9 then allsmears=3
label allsmears "Number of smears"

cls
* Count all positive quantified smears
* (include scanty not quantified)
gen i allpos1=0
if result1>0 and result1<4 then allpos1=1
if result1=5               then allpos1=1
gen i allpos2=0
if result2>0 and result2<4 then allpos2=1
if result2=5               then allpos2=1
gen i allpos3=0
if result3>0 and result3<4 then allpos3=1
if result3=5               then allpos3=1
gen i allpos=allpos1+allpos2+allpos3
label allpos "Number of positive smears"

cls
* Count all scanty smears
* (include scanty not quantified)
gen i scantpos1=0
* (include scanty not quantified)
if result1>0 and result1<1 then scantpos1=1
if result1=5               then scantpos1=1
gen i scantpos2=0
if result2>0 and result2<1 then scantpos2=1
if result2=5               then scantpos2=1
gen i scantpos3=0
if result3>0 and result3<1 then scantpos3=1
if result3=5               then scantpos3=1
gen i scantypos=scantpos1+scantpos2+scantpos3
label scantypos "Number of scanty positive smears"

cls
* Count all low positive smears
* (include scanty not quantified)
gen i lowpos1=0
if result1>0 and result1<2 then lowpos1=1
if result1=5               then lowpos1=1
gen i lowpos2=0
if result2>0 and result2<2 then lowpos2=1
if result2=5               then lowpos2=1
gen i lowpos3=0
if result3>0 and result3<2 then lowpos3=1
if result3=5               then lowpos3=1
gen i lowpos=lowpos1+lowpos2+lowpos3
label lowpos "Number of low positive smears"

keep regyear mmseq allsmears allpos scantypos lowpos
savedata "temp_02.rec" /replace
**************************************************
* 3) Aggregate data

cls
close
read "temp_02.rec"

agg mmseq /sum=allsmears /sum=allpos /sum=scantypos /sum=lowpos /close
drop n nallsm1 nallpos nscantpos nlowpos
rename sumallsm1 to allsmears
rename sumallpos to allpos
rename sumscant1 to scantpos
rename sumlowpos to lowpos
```

```
drop n nallsm1 nallpos nscant1 nlowpos
savedata "temp_03.rec" /replace
************************************************
* 4) Make SPC charts

set option spc= /sizex=500 /sizey=350
set graph font size=9

cls
close
logclose

read "temp_03.rec"

set echo=off
pchart allpos allsmears mmseq /xtext="Month of observation" /bw \
       /ytext="Per cent low positive" \
       /ti="Proportion positive smears among all smears" \
       /sub="Uganda, 1999-2001" \
       /fn="Course Part D, Exercise D_EX02, in EpiData Analysis" \
       /b=12 /b=24 \
       /xlined=12.5 /xlined=24.5 \
       /t1 \
       /text="120,60,1999,0" \
       /text="260,60,2000,0" \
       /text="400,60,2001,0"

pchart scantpos allsmears mmseq /xtext="Month of observation" /bw \
       /ytext="Per cent scanty positive" \
       /ti="Proportion scanty pos smears among all smears" \
       /sub="Uganda, 1999-2001" \
       /fn="Course Part D, Exercise D_EX02, in EpiData Analysis" \
       /b=12 /b=24 \
       /xlined=12.5 /xlined=24.5 \
       /t1 \
       /text="120,60,1999,0" \
       /text="250,60,2000,0" \
       /text="400,60,2001,0"

pchart lowpos allsmears mmseq /xtext="Month of observation" /bw \
       /ytext="Per cent scanty positive" \
       /ti="Proportion low positive smears among all smears" \
       /sub="Uganda, 1999-2001" \
       /fn="Course Part D, Exercise D_EX02, in EpiData Analysis" \
       /b=12 /b=24 \
       /xlined=12.5 /xlined=24.5 \
       /t1 \
       /text="120,60,1999,0" \
       /text="260,60,2000,0" \
       /text="400,60,2001,0"

pchart lowpos allpos mmseq /xtext="Month of observation" /bw \
       /ytext="Per cent low positive" \
       /ti="Proportion low-pos smears among all positive smears" \
       /sub="Uganda, 1999-2001" \
       /fn="Course Part D, Exercise D_EX02, in EpiData Analysis" \
       /b=12 /b=24 \
       /xlined=12.5 /xlined=24.5 \
       /t1 \
       /text="130,60,1999,0" \
       /text="260,60,2000,0" \
       /text="400,60,2001,0"

pchart scantpos allpos mmseq /xtext="Month of observation" /bw \
       /ytext="Per cent scanty positive" \
       /ti="Proportion scanty pos smears among all positive smears" \
       /sub="Uganda, 1999-2001" \
       /fn="Course Part D, Exercise D_EX02, in EpiData Analysis" \
       /b=12 /b=24 \
       /xlined=12.5 /xlined=24.5 \
       /t1 \
       /text="120,60,1999,0" \
       /text="250,60,2000,0" \
       /text="400,60,2001,0"

set echo=on
```