Homework 10 – Tuo Wang

1, Increase the variance of the errors and the value of alpha. What do you find about bias? Median bias?

I increased the alpha from 200 to 600 and the variance of the errors from 5 to 8. Before when alpha was 200 and the variance was 5 there was bias but there was barely any bias in the median if at all none. Now there is more median bias before it was 3.02 for linear. Now there is a bias of 0.2 since the estimate is 2.8240 in the functional form of linear but there is no bias when it is log 300 for the median. In the long form we even see a reduction of bias for the median since it was 3.02. However, with the bias we find that it is even greater than before with bot the linear and the log transform.

dim bbox 5 1000 lbbox 5 1000

set nodoecho

do#=1,1000

smpl 1 100

genr x=uni(10)

genr e=nor(1)

genr y=600+3\*x+nor(8)

genr ey=exp(y)

genr ex=exp(x)

?box y x/all coef=bbox:#

?box ey ex/all coef=lbbox:#

endo

matrix bbox=bbox’

matrix lbbox=lbbox’

smpl 1 1000

stat bbox lbbox/pmedian

…NOTE…TREATING COLUMNS OF BBOX AS VECTORS

NAME N MEAN ST.DEV VARIANCE MINIMUM MAXIMUM

BBOX 1000 41.934 791.08 0.6258E 0.5622E 24923.

BBOX 1000 70745. 0.199E+07 0.3976E+13 15.818 0.6E+08

BBOX 1000 1.0252 0.31957 0.10213 0.25000 2.9700

BBOX 1000 1.0000 0.000E+00 0.000E+00 1.0000 1.0000

BBOX 1000 1.0252 0.31957 0.10213 0.25000 2.9700

…NOTE…TREATING COLUMNS OF BBOX AS VECTORS

LBBOX 1000 7600.6 63468. 0.40282E+10 0.12818E-04 0.60251E+06

LBBOX 1000 0.11751E+06 0.93282E+06 0.87014E+12 50.000 0.87347E+07

LBBOX 1000 0.50000E-04 0.77043E-02 0.59357E-04 -0.20000E-010.20000E-01

LBBOX 1000 1.0000 0.00000E+00 0.00000E+00 1.0000 1.0000

LBBOX 1000 0.50000E-04 0.77043E-02 0.59357E-04 0.20000E-02 0.20000E-01

VARIABLE = BBOX

MEDIAN = 2.8240

LOWER 25% = 1.0991 UPPER25% = 9.0241 INTERQUARTILE RANGE = 6.925

MODE NOT APPLICABLE

VARIABLE = BBOX

MEDIAN = 570.57

LOWER 25% = 208.00 UPPER25% = 1802.7 INTERQUARTILE RANGE = 1595.

MODE NOT APPLICABLE

VARIABLE = BBOX

MEDIAN = 0.99000

LOWER 25% = 0.80000 UPPER25% = 1.2000 INTERQUARTILE RANGE = 0.4000

MODE=0.72000 WITH 21 OBSERVATIONS

VARIABLE = BBOX

MEDIAN = 1.0000

LOWER 25% = 1.0000 UPPER25% = 1.0000 INTERQUARTILE RANGE = 0.0000E+00

MODE = 1.0000 WITH 1000 OBSERVATIONS

VARIABLE = BBOX

MEDIAN = 0.99000

LOWER 25% = 0.80000 UPPER25% = 1.2000 INTERQUARTILE RANGE = 0.4000

MODE=0.72000 WITH 21 OBSERVATIONS

VARIABLE = LBBOX

MEDIAN = 3.0014

LOWER 25% = 0.76987E-02 UPPER25% = 1123.1 INTERQUARTILE RANGE = 1123.

MODE NOT APPLICABLE

VARIABLE = LBBOX

MEDIAN = 600.11

LOWER 25% = 99.759 UPPER25% = 39054 INTERQUARTILE RANGE = 0.3895E+05

MODE NOT APPLICABLE

VARIABLE = LBBOX

MEDIAN = 0.00000E+00

LOWER 25% = -0.10000E+00 UPPER25% = 0.10000E-01 INTERQUARTILE RANGE =0.2000E-01

MODE = 0.00000E+00 WITH 470 OBSERVATIONS

VARIABLE = LBBOX

MEDIAN = 1.0000

LOWER 25% = 1.000 UPPER25% = 1.0000 INTERQUARTILE RANGE = 0.0000E+00

MODE = 1.0000 WITH 1000 OBSERVATIONS

VARIABLE = LBBOX

MEDIAN = 0.00000E+00

LOWER 25% = -0.100000E-01 UPPER25% = 0.10000E-01 INTERQUARTILE RANGE = 0.2000E-01

MODE = 0.00000E+00 WITH 470 OBSERVATIONS

2. Drop the sample from 100. Bias/median bias does what?

I dropped the sample size from 100 to 50. We notice that the bias is increases significantly and the median bias does increase a slight bit since it goes from 3.1000 to 3.1724 for the linear form and it remains slightly unbiased in the long form is it is 3.03

dim bbox 5 1000 lbbox 5 1000

set nodoecho

do#=1,1000

smpl 1 50

genr x=uni(10)

genr e=nor(1)

genr y=200+3\*x+nor(5)

genr ey=exp(y)

genr ex=exp(x)

?box y x/all coef=bbox:#

?box ey ex/all coef=lbbox:#

endo

matrix bbox=bbox’

matrix lbbox=lbbox’

smpl 1 1000

stat bbox lbbox/pmedian

…NOTE…TREATING COLUMNS OF BBOX AS VECTORS

NAME N MEAN ST.DEV VARIANCE MINIMUM MAXIMUM

BBOX 1000 7.5926 38.222 1468.6 0.20911 1162.4

BBOX 1000 1145.3 16628. 0.2754E+09 16.620 0.5E+06

BBOX 1000 1.0325 0.28513 0.812E-01 0.37000 2.6500

BBOX 1000 1.0000 0.000E+00 0.000E+00 1.0000 1.0000

BBOX 1000 1.0325 0.28513 0.81296E-01 0.37000 2.6500

…NOTE…TREATING COLUMNS OF BBOX AS VECTORS

LBBOX 1000 35.971 158.35 25074 0.64213E 1830.7

LBBOX 1000 564.43 1427.8 0.2038E+09 24.992 14065.

LBBOX 1000 -0.15E 0.1124E 0.1264E-01 0.4000 0.3000E

LBBOX 1000 1.0000 0.000E+00 0.000E+00 1.0000 1.0000

LBBOX 1000 -0.15E 0.1124E 0.1264E-01 0.4000 0.3000E

VARIABLE = BBOX

MEDIAN = 3.1724

LOWER 25% = 1.4942 UPPER25% = 6.6744 INTERQUARTILE RANGE = 5.280

MODE NOT APPLICABLE

VARIABLE = BBOX

MEDIAN = 218.65

LOWER 25% = 97.092 UPPER25% = 488.77 INTERQUARTILE RANGE = 391.7

MODE NOT APPLICABLE

VARIABLE = BBOX

MEDIAN = 1.0200

LOWER 25% = 0.8300 UPPER25% = 1.2000 INTERQUARTILE RANGE = 0.370

..WARNING..MULTIPLE MODE – THE MAXIMUM IS REPORTED

VARIABLE = BBOX

MEDIAN = 1.0000

LOWER 25% = 1.0000 UPPER25% = 1.0000 INTERQUARTILE RANGE = 0.000E

MODE = 1.0000 WITH 1000 OBSERVATIONS

VARIABLE = BBOX

MEDIAN = 1.0200

LOWER 25% = 0.8300 UPPER25% = 1.2000 INTERQUARTILE RANGE = 0.370

..WARNING..MULTIPLE MODE – THE MAXIMUM IS REPORTED

VARIABLE = LBBOX

MEDIAN = 3.0310

LOWER 25% = 0.37790 UPPER25% = 23.609 INTERQUARTILE RANGE = 23.23

MODE NOT APPLICABLE

VARIABLE = LBBOX

MEDIAN = 199.65

LOWER 25% = 86.577 UPPER25% = 631.62 INTERQUARTILE RANGE = 545.0

MODE NOT APPLICABLE

VARIABLE = LBBOX

MEDIAN = 0.00000E+00

LOWER 25% = -0.100 UPPER25% = 0.1000 INTERQUARTILE RANGE = 0.200E

MODE = 0.0000E+00 WITH 294 OBSERVATIONS

VARIABLE = LBBOX

MEDIAN = 1.0000

LOWER 25% = 1.0000 UPPER25% = 1.0000 INTERQUARTILE RANGE = 0.000E

MODE = 1.0000 WITH 1000 OBSERVATIONS

VARIABLE = LBBOX

MEDIAN = 0.00000E+00

LOWER 25% = -0.100 UPPER25% = 0.1000 INTERQUARTILE RANGE = 0.200E

MODE = 0.0000E+00 WITH 294 OBSERVATIONS