2.4.1: SAS - Simultaneous Inference and Regression Through Origin

Dr. Bean – Stat 5100

```
/*********************************************************/

/* Input Toluca data (recall Ch. 1 example) */
data toluca; input lotsize workhours @@; cards;
 80 399   30  121   50  221   90  376   70  361   60  224
120 546   80  352  100  353   50  157   40  160   70  252
 90 389   20  113  110  435  100  420   30  212   50  268
 90 377  110  421   30  273   90  468   40  244   80  342
 70 323
  ;
run;

/* Simultaneous 95% interval estimation of betas  */
proc reg data=toluca;
  model workhours = lotsize / clb alpha=.025;
  title1 'Simultaneous 95% confidence intervals on betas';
run;
```

Simultaneous 95% confidence intervals on betas

Parameter Estimates

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	97.5% Confidence Limits	
Intercept	1	62.36586	26.17743	2.38	0.0259	-0.40436	125.13607
lotsize	1	3.57020	0.34697	10.29	<.0001	2.73821	4.40220

```
/* Simultaneous 90% interval estimation of mean workhours
   at lotsize levels 30, 65, 100  (using Working-Hotelling
   and Bonferroni)
 */
data dummy; input lotsize check; cards;
  30  1
  65  1
  100 1
  ;
data temp; set toluca dummy;
proc reg data=temp noprint;
  model workhours = lotsize;
  output out=out1 p=Yhat stdp=seYhat;
    /* KEY: stdp is SE of mean prediction */
data out1; set out1;
  alpha = 0.10;  /* 1-alpha is simult. conf. level */
  p = 2;         /* # of beta's (including intercept) */
  n = 25;        /* sample size */
  g = 3;         /* number of simultaneous intervals */
  W = sqrt(p*finv(1-alpha,p,n-p));  /* WH crit. val. */
  t = tinv(1-alpha/(2*g),n-p);      /* Bonf. crit. val. */
  WH_upper = Yhat + W*seYhat;
  WH_lower = Yhat - W*seYhat;
  B_upper = Yhat + t*seYhat;
  B_lower = Yhat - t*seYhat;
proc print data=out1;
  where check = 1;
  var lotsize Yhat seYhat WH_lower WH_upper
      B_lower B_upper;
  title1
   'Simultaneous 90% interval estimation of mean response';
  title2
   'at three X-levels, using Working-Hotelling and
    Bonferroni';
run;
```

Simultaneous 90% interval estimation of mean response
at three X-levels, using Working-Hotelling and Bonferroni

Obs	lotsize	Yhat	seYhat	WH_lower	WH_upper	B_lower	B_upper
26	30	169.472	16.9697	131.154	207.790	131.057	207.887

27	65	294.429	9.9176	272.035	316.823	271.978	316.880
28	100	419.386	14.2723	387.159	451.613	387.077	451.695

```
/* Simultaneous 95% prediction limits on next two lots,
   with sizes 80 and 100 units (using Scheffe and
   Bonferroni)
 */
data dummy; input lotsize check; cards;
  80  1
  100 1
  ;
data temp; set toluca dummy;
proc reg data=temp noprint;
  model workhours = lotsize;
  output out=out1 p=Yhat stdi=seYhatnew;
    /* KEY: stdi is SE of individual prediction */
data out1; set out1;
  alpha = 0.05;  /* 1-alpha is simult. pred. level */
  p = 2;         /* # of beta's (including intercept) */
  n = 25;        /* sample size */
  g = 2;         /* number of simultaneous intervals */
  S = sqrt(g*finv(1-alpha,g,n-p));  /* Scheffe crit val */
  t = tinv(1-alpha/(2*g),n-p);      /* Bonf. crit. val. */
  S_upper = Yhat + S*seYhatnew;
  S_lower = Yhat - S*seYhatnew;
  B_upper = Yhat + t*seYhatnew;
  B_lower = Yhat - t*seYhatnew;
proc print data=out1;
  where check = 1;
  var lotsize Yhat seYhatnew S_lower S_upper
      B_lower B_upper;
  title1 'Simultaneous 95% interval estimation of
          individual prediction';
  title2 'at two X-levels, using Scheffe and Bonferroni';
run;
```

Simultaneous 95% interval estimation of individual prediction at two X-levels, using Scheffe and Bonferroni

Obs	lotsize	Yhat	seYhatnew	S_lower	S_upper	B_lower	B_upper
26	80	347.982	49.9110	217.407	478.557	228.302	467.662
27	100	419.386	50.8666	286.311	552.461	297.414	541.358

```
/*********************************************************/

/* Regression through origin example: plumbing supplies
   company looking at relationship between number of
   work units (X) and labor costs (Y) at its 12 warehouses
 */

data warehouse; input work cost @@; cards;
 20 114   196 921   115 560    50 245   122 575   100 475
 33 138   154 727    80 375   147 670   182 828   160 762
  0   .
   ;
proc reg data=warehouse;
  model cost = work / noint;
  output out=out1 p=pred;
  title1 'Regression through origin';
run;
```

Regression through origin

Number of Observations Read	13
Number of Observations Used	12
Number of Observations with Missing Values	1

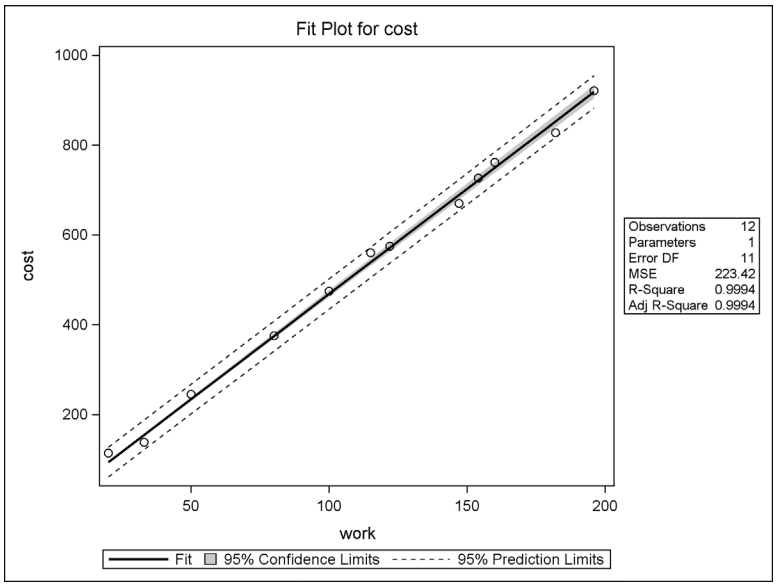
Note: No intercept in model. R-Square is redefined.

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	4191980	4191980	18762.5	<.0001
Error	11	2457.65933	223.42358		
Uncorrected Total	12	4194438			

Root MSE	14.94736	**R-Square**	0.9994
Dependent Mean	532.50000	**Adj R-Sq**	0.9994
Coeff Var	2.80702		

Parameter Estimates					
Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|
work	1	4.68527	0.03421	136.98	<.0001




```
proc sort data=out1; by work;
proc sgplot data=out1;
  scatter x=work y=cost /
        markerattrs=(symbol=CIRCLEFILLED size=2pt);
  series x=work y=pred / lineattrs=(pattern=solid);
  xaxis values=(0 to 200 by 50);
  yaxis values=
       (0 to 1000 by 200);
run;
/* Note forced inclusion of
  work=0 dummy observation
   for graphical purposes */
```