ผนวก ข

ข้อมูลที่ใช้ในการวิจัย

File Name: PiDigits.dat

Dataset Name: PiDigits

Description: This is a constructed/fabricated data set to test accuracy in summary statistic calculations.The numbers are the first 5000 digits of the mathematical constant pi (= 3.1415926535897932384...).

Stat Category: Univariate

Reference: Mathematics of Computation. January 1962, page 76.

Data: Constructed Variable

--> 1 Response : y = pi digits

--> 0 Predictors

--> 5000 Observations

Model: Lower Level of Difficulty

--> 2 Parameters : mu, sigma

--> 1 Response Variable : y

--> 0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 4.53480000000000

Sample Standard Deviation (denom. = n-1) s: 2.86733906028871

Sample Autocorrelation Coefficient (lag 1) r(1): -0.00355099287237972

Number of Observations: 5000

File Name: Lottery.dat

Dataset Name: Lottery

Description: This is an observed/"real world" data set consisting of 218 lottery values from September 3, 1989 to April 14, 1990 (32 weeks).One 3-digit random number (from 000 to 999) is drawn per day, 7 days per week for most weeks, but fewer days per week for some weeks.We here use this data to test accuracy in summary statistics calculations.

Stat Category: Univariate: Summary Statistics

Reference: None

Data: "Real World"

1 Response : y = 3-digit random number

0 Predictors

218 Observations

Model: Lower Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 518.958715596330

Sample Standard Deviation (denom. = n-1) s: 291.699727470969

Sample Autocorrelation Coefficient (lag 1) r(1): -0.120948622967393

Number of Observations: 218

File Name: Lew.dat

Dataset Name: Lew (Beam Deflection Data)

Description: This is an observed/"real world" data set consisting of 200 deflections of a steel-concrete beam while subjected to periodic pressure.The experimenter was H. S. Lew of the Center for Building Technology at NIST. We here use this data to test accuracy in summary statistics calculations.

Stat Category: Univariate: Summary Statistics

Reference: None

Data: "Real World"

1 Response : y = beam deflection

0 Predictors

200 Observations

Model: Lower Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: -177.435000000000

Sample Standard Deviation (denom. = n-1) s: 277.332168044316

Sample Autocorrelation Coefficient (lag 1) r(1): -0.307304800605679

Number of Observations: 200

File Name: Mavro.dat

Dataset Name: Mavro (Filter Transmittance Data)

Description: This is an observed/"real world" data set consisting of 50 transmittance measurements (at a sampling rate of 10 observations per second) from a filter with a nominal value of 2. The experimenter was Radu Mavrodineaunu, a member of the chemistry staff at NIST. We here use this data to test accuracy in summary statistics calculations.

Stat Category: Univariate: Summary Statistics

Reference: None

Data: "Real World"

1 Response : y = transmittance

0 Predictors

50 Observations

Model: Lower Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 2.00185600000000

Sample Standard Deviation (denom. = n-1) s: 0.000429123454003053

Sample Autocorrelation Coefficient (lag 1) r(1): 0.937989183438248

Number of Observations: 50

File Name: Michelso.dat

Dataset Name: Michelso (Speed of Light Data, in millions of meters per second)

Description: This is an observed/"real world" data set consisting of 100 measurements of the speed of light in air. This classic experiment was carried out by Michelson is 1879. We here use this data to test accuracy in summary statistics calculations.

Stat Category: Univariate: Summary Statistics

Reference: Dorsey, Ernest N. (1944). The Velocity of Light. Transactions of the American Philiosophical Society, Volume 34, Part 1, Pages 1-110, Table 22.

Model: Lower Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 299.852400000000

Sample Standard Deviation (denom. = n-1) s: 0.0790105478190518

Sample Autocorrelation Coefficient (lag 1) r(1): 0.535199668621283

Number of Observations: 100

File Name: NumAcc1.dat

Dataset Name: NumAcc1

Description: This is a constructed/fabricated data set to test accuracy in summary statistic calculations.The numbers are large (8-digit integers) and differ only in the last decimal place.

Note--by construction, this data set has

sample mean = 10000002 (exact)

sample standard deviation = 1 (exact)

sample autocorrelation coef. = -0.5 (exact)

Stat Category: Univariate: Summary Statistics

Reference: Simon, Stephen D. and Lesage, James P. (1989). Assessing the Accuracy of ANOVA Caluclations in Statistical Software", Computational Statistics & data Analysis, 8, pp. 325-332.

Data: Constructed

1 Response : y

0 Predictors

3 Observations

Model: Lower Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 10000002 (exact)

Sample Standard Deviation (denom. = n-1) s: 1 (exact)

Sample Autocorrelation Coefficient (lag 1) r(1): -0.5 (exact)

Number of Observations: 3

File Name: NumAcc2.dat

Dataset Name: NumAcc2

Description: This is a constructed/fabricated data set to test accuracy in summary statistic calculations. The numbers are 2-digit floating point values and differ only in the last decimal place.

Note--by construction, this data set has

sample mean = 1.2 (exact)

sample standard deviation = 0.1 (exact)

sample autocorrelation coef. = -0.999 (exact)

Stat Category: Univariate

Reference: Simon, Stephen D. and Lesage, James P. (1989).Assessing the Accuracy of ANOVA Caluclations in Statistical Software", Computational Statistics & data Analysis, 8, pp. 325-332.

Data: Constructed

1 Response : y

0 Predictors

1001 Observations

Model: Average Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 1.2 (exact)

Sample Standard Deviation (denom. = n-1) s: 0.1 (exact)

Sample Autocorrelation Coefficient (lag 1) r(1): -0.999 (exact)

Number of Observations: 1001

File Name: NumAcc3.dat

Dataset Name: NumAcc3

Description: This is a constructed/fabricated data set to test accuracy in summary statistic calculations.The numbers are 8-digit floating point values and differ only in the last decimal place.

Note--by construction, this data set has

sample mean = 1000000.2 (exact)

sample standard dev. = 0.1 (exact)

sample autocorr. coef. = -0.999 (exact)

Stat Category: Univariate: Summary Statistics

Reference: Simon, Stephen D. and Lesage, James P. (1989). Assessing the Accuracy of ANOVA Caluclations in Statistical Software", Computational Statistics & data Analysis, 8, pp. 325-332.

Data: Constructed

1 Response : y

0 Predictors

1001 Observations

Model: Average Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 1000000.2 (exact)

Sample Standard Deviation (denom. = n-1) s: 0.1 (exact)

Sample Autocorrelation Coefficient (lag 1) r(1): -0.999 (exact)

Number of Observations: 1001

File Name: NumAcc4.dat

Dataset Name: NumAcc4

Description: This is a constructed/fabricated data set to test accuracy in summary statistic calculations.The numbers are 9-digit floating point values and differ only in the last decimal place.

sample mean = 10000000.2 (exact)

sample standard dev. = 0.1 (exact)

sample autocorr. coef. = -0.999 (exact)

Stat Category: Univariate

Reference: Simon, Stephen D. and Lesage, James P. (1989).Assessing the Accuracy of ANOVA Caluclations in Statistical Software", Computational Statistics & data Analysis, 8, pp. 325-332.

Data: Constructed

1 Response : y

0 Predictors

1001 Observations

Model: Higher Level of Difficulty

2 Parameters : mu, sigma

1 Response Variable : y

0 Predictor Variables

y = mu + e

Certified Values

Sample Mean ybar: 10000000.2 (exact)

Sample Standard Deviation (denom. = n-1) s: 0.1 (exact)

Sample Autocorrelation Coefficient (lag 1) r(1): -0.999 (exact)

Number of Observations: 1001

Dataset Name: SmLs01 (SmLs01.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
21 Replicates/Cell
189 Observations
1 Constant Leading Digit
Lower Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.68000000000000E+00	2.10000000000000E-01
Within Treatment	180	1.80000000000000E+00	1.00000000000000E-02
F Statistic		2.10000000000000E+01	

Dataset Name: SmLs02 (SmLs02.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
201 Replicates/Cell
1809 Observations
1 Constant Leading Digit
Lower Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.60800000000000E+01	2.01000000000000E+00
Within Treatment	1800	1.80000000000000E+01	1.00000000000000E-02

F Statistic 2.01000000000000E+02

Dataset Name: SmLs03 (SmLs03.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
2001 Replicates/Cell
18009 Observations
1 Constant Leading Digit
Lower Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.60080000000000E+02	2.00100000000000E+01
Within Treatment	18000	1.80000000000000E+02	1.00000000000000E-02
F Statistic		2.00100000000000E+03	

Dataset Name: SmLs04 (SmLs04.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
21 Replicates/Cell
189 Observations
7 Constant Leading Digits
Average Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.68000000000000E+00	2.10000000000000E-01
Within Treatment	180	1.80000000000000E+00	1.00000000000000E-02
F Statistic		2.10000000000000E+01	

Dataset Name: SmLs05 (SmLs05.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
201 Replicates/Cell
1809 Observations
7 Constant Leading Digits
Average Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.60800000000000E+01	2.01000000000000E+00
Within Treatment	1800	1.80000000000000E+01	1.00000000000000E-02
F Statistic		2.01000000000000E+02	

Dataset Name: SmLs06 (SmLs06.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
2001 Replicates/Cell
18009 Observations
7 Constant Leading Digits
Average Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.60080000000000E+02	2.00100000000000E+01
Within Treatment	18000	1.80000000000000E+02	1.00000000000000E-02
F Statistic		2.00100000000000E+03	

Dataset Name: SmLs07 (SmLs07.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
21 Replicates/Cell
189 Observations
13 Constant Leading Digits
Higher Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.68000000000000E+00	2.10000000000000E-01
Within Treatment	180	1.80000000000000E+00	1.00000000000000E-02
F Statistic		2.10000000000000E+01	

Dataset Name: SmLs08 (SmLs08.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
201 Replicates/Cell
1809 Observations
13 Constant Leading Digits
Higher Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.60800000000000E+01	2.01000000000000E+00
Within Treatment	1800	1.80000000000000E+01	1.00000000000000E-02
F Statistic		2.01000000000000E+02	

Dataset Name: SmLs09 (SmLs09.dat)

Procedure: Analysis of Variance

Reference: Simon, Stephen D. and Lesage, James P. (1989). "Assessing the Accuracy of ANOVA Calculations in Statistical Software".Computational Statistics & Data Analysis, 8, pp. 325-332.

Data: 1 Factor
9 Treatments
2001 Replicates/Cell
18009 Observations
13 Constant Leading Digits
Higher Level of Difficulty
Generated Data

Model: 10 Parameters (mu,tau_1, ... , tau_9)
y_{ij} = mu + tau_i + epsilon_{ij}

Certified Values:

Source of Variation	df	Sums of Squares	Mean Squares
Between Treatment	8	1.60080000000000E+02	2.00100000000000E+01
Within Treatment	18000	1.80000000000000E+02	1.00000000000000E-02
F Statistic		2.00100000000000E+03	

Dataset Name: Norris (Norris.dat)

Procedure: Linear Least Squares Regression

Reference: Norris, J., NIST. Calibration of Ozone Monitors.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
36 Observations
Lower Level of Difficulty
Observed Data

Model: Linear Class 2 Parameters (B0,B1)
y = B0 + B1*x + e

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B0	-0.262323073774029	0.232818234301152
B1	1.00211681802045	0.429796848199937E-03

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	1	4255954.13232369	4255954.13232369
Residual	34	26.6173985294224	0.782864662630069

F Statistic 5436385.54079785

Dataset Name: NoInt1 (NoInt1.dat)

Procedure: Linear Least Squares Regression

Reference: Eberhardt, K., NIST.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
11 Observations
Average Level of Difficulty
Generated Data

Model: Linear Class 1 Parameter (B1)
y = B1*x + e

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B1	2.07438016528926	0.165289256198347E-01

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	1	200457.727272727	200457.727272727
Residual	10	127.272727272727	12.7272727272727

F Statistic 15750.2500000000

Dataset Name: NoInt2 (NoInt2.dat)

Procedure: Linear Least Squares Regression

Reference: Eberhardt, K., NIST.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
3 Observations
Average Level of Difficulty
Generated Data

Model: Linear Class 1 Parameter (B1)
y = B1*x + e

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B1	0.727272727272727	0.420827318078432E-01

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	1	40.7272727272727	40.7272727272727
Residual	2	0.272727272727273	0.136363636363636

F Statistic 298.6666666666667

Dataset Name: Wampler1 (Wampler1.dat)

Procedure: Linear Least Squares Regression

Reference: Wampler, R. H. (1970). A Report of the Accuracy of Some Widely-Used Least Squares Computer Programs. Journal of the American Statistical Association, 65, pp. 549-565.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
21 Observations
Higher Level of Difficulty
Generated Data

Model: Polynomial Class 6 Parameters (B0,B1,...,B5)

y = B0 + B1*x + B2*(x**2) + B3*(x**3)+ B4*(x**4) + B5*(x**5)

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B0	1.00000000000000	0.000000000000000
B1	1.00000000000000	0.000000000000000
B2	1.00000000000000	0.000000000000000
B3	1.00000000000000	0.000000000000000
B4	1.00000000000000	0.000000000000000
B5	1.00000000000000	0.000000000000000

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	5	18814317208116.7	3762863441623.33
Residual	15	0.000000000000000	0.000000000000000
F Statistic		Infinity	

Dataset Name: Wampler2 (Wampler2.dat)

Procedure: Linear Least Squares Regression

Reference: Wampler, R. H. (1970). A Report of the Accuracy of Some Widely-Used Least Squares Computer Programs. Journal of the American Statistical Association, 65, pp. 549-565.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
21 Observations
Higher Level of Difficulty
Generated Data

Model: Polynomial Class 6 Parameters (B0,B1,...,B5)
y = B0 + B1*x + B2*(x**2) + B3*(x**3)+ B4*(x**4) + B5*(x**5)

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B0	1.00000000000000	0.000000000000000
B1	0.100000000000000	0.000000000000000
B2	0.100000000000000E-01	0.000000000000000
B3	0.100000000000000E-02	0.000000000000000
B4	0.100000000000000E-03	0.000000000000000
B5	0.100000000000000E-04	0.000000000000000

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	5	6602.91858365167	1320.58371673033
Residual	15	0.000000000000000	0.000000000000000
F Statistic		Infinity	

Dataset Name: Wampler3 (Wampler3.dat)

Procedure: Linear Least Squares Regression

Reference: Wampler, R. H. (1970). A Report of the Accuracy of Some Widely-Used Least Squares Computer Programs. Journal of the American Statistical Association, 65, pp. 549-565.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
21 Observations
Higher Level of Difficulty
Generated Data

Model: Polynomial Class 6 Parameters (B0,B1,...,B5)

y = B0 + B1*x + B2*(x**2) + B3*(x**3)+ B4*(x**4) + B5*(x**5)

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B0	1.00000000000000	2152.32624678170
B1	1.00000000000000	2363.55173469681
B2	1.00000000000000	779.343524331583
B3	1.00000000000000	101.475507550350
B4	1.00000000000000	5.64566512170752
B5	1.00000000000000	0.112324854679312

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	5	18814317208116.7	3762863441623.33
Residual	15	83554268.0000000	5570284.53333333
F Statistic		675524.458240122	

Dataset Name: Wampler4 (Wampler4.dat)

Procedure: Linear Least Squares Regression

Reference: Wampler, R. H. (1970). A Report of the Accuracy of Some Widely-Used Least Squares Computer Programs.Journal of the American Statistical Association, 65, pp. 549-565.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
21 Observations
Higher Level of Difficulty
Generated Data

Model: Polynomial Class 6 Parameters (B0,B1,...,B5)
y = B0 + B1*x + B2*(x**2) + B3*(x**3)+ B4*(x**4) + B5*(x**5)

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B0	1.00000000000000	215232.624678170
B1	1.00000000000000	236355.173469681
B2	1.00000000000000	77934.3524331583
B3	1.00000000000000	10147.5507550350
B4	1.00000000000000	564.566512170752
B5	1.00000000000000	11.2324854679312

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	5	18814317208116.7	3762863441623.33
Residual	15	835542680000.000	55702845333.3333
F Statistic		67.5524458240122	

Dataset Name: Wampler5 (Wampler5.dat)

Procedure: Linear Least Squares Regression

Reference: Wampler, R. H. (1970). A Report of the Accuracy of Some Widely-Used Least Squares Computer Programs.Journal of the American Statistical Association, 65, pp. 549-565.

Data: 1 Response Variable (y)
1 Predictor Variable (x)
21 Observations
Higher Level of Difficulty
Generated Data

Model: Polynomial Class 6 Parameters (B0,B1,...,B5)

y = B0 + B1*x + B2*(x**2) + B3*(x**3)+ B4*(x**4) + B5*(x**5)

Certified Regression Statistics

Parameter	Estimate	Standard Deviation of Estimate
B0	1.00000000000000	21523262.4678170
B1	1.00000000000000	23635517.3469681
B2	1.00000000000000	7793435.24331583
B3	1.00000000000000	1014755.07550350
B4	1.00000000000000	56456.6512170752
B5	1.00000000000000	1123.24854679312

Certified Analysis of Variance Table

Source of Variation	df	Sums of Squares	Mean Squares
Regression	5	18814317208116.7	3762863441623.33
Residual	15	0.835542680000000E+16	557028453333333.
F Statistic		6.7552445824012241E-03	