Szelcsanyi Boris 19-606-656
Source

```
new;
cls;
rndseed 1234;

// Load data
file = "michelson.dat";
fh = dataopen(file, "read");
X = loadd(file);

X_names = getHeaders(file);
X_names;

// Task 1
nrows = rows(X);
nrows;

dstatmt(X);
// Looking at the output of dstatmt(x) we can conclude that there are max. 5 with 20 run experiments which equal to 100 observations
print "_____\n";

// Task 2
// indexing/slicing data
sub_sample_e1 = X;
sub_sample_e1 = selif(sub_sample_e1, sub_sample_e1[., "Expt"] .== 1);
sub_sample_e1 = sub_sample_e1[.,"Speed"];
speed_e1 = sub_sample_e1[.,"Speed"];

sub_sample_e5 = X;
sub_sample_e5 = selif(sub_sample_e5, sub_sample_e5[., "Expt"] .== 5);
sub_sample_e5_ = sub_sample_e5[.,"Speed"];
speed_e5 = sub_sample_e5[.,"Speed"];

// procedure for variance
proc(1) = variance(data);
    local mu, sigma_2, size_n;
    mu = meanc(data);
    size_n = rows(data);
    sigma_2 = 0;
    for i (1, rows(data), 1);
        sigma_2 = sigma_2 + ((data[i] - mu)^2);
    endfor;
    retp(sigma_2/size_n);
endp;

print "Variance for Speed in Expt 1 is " variance(speed_e1);
print "Variance for Speed in Expt 5 is " variance(speed_e5);


// Task 3
print "_____\n";
sigma_2 = 7500;
mu1 = meanc(speed_e1);
mu2 = meanc(speed_e5);
size_e1_e2 = 20;
t = ( sqrt(size_e1_e2) * (mu1 - mu2) ) / sqrt(sigma_2 + sigma_2);
t2 = ( sqrt(size_e1_e2) * (mu1 - mu2) ) / sqrt(variance(speed_e1) + variance(speed_e5));
print "t test for both sub samples, where s1 = s2 = 7500 :" t;
print "t test for both sub samples, where s1 != s2 using results from procedure: " t2;
print "Since abs(t) > 2.05, for both approaches, it follows means are different";

print "_____\n";
// Task 4

file2 = "michelson_binary.dat";
```

```
48
49  // Task 3
50  print "_____\n";
51  sigma_2 = 7500;
52  mu1 = meanc(speed_e1);
53  mu2 = meanc(speed_e5);
54  size_e1_e2 = 20;
55  t = ( sqrt(size_e1_e2) * (mu1 - mu2) ) / sqrt(sigma_2 + sigma_2);
56  t2 = ( sqrt(size_e1_e2) * (mu1 - mu2) ) / sqrt(variance(speed_e1) + variance(speed_e5));
57  print "t test for both sub samples, where s1 = s2 = 7500 :" t;
58  print "t test for both sub samples, where s1 != s2 using results from procedure: " t2;
59  print "Since abs(t) > 2.05, for both approaches, it follows means are different";
60
61  print "_____\n";
62  // Task 4
63
64  file2 = "michelson_binary.dat";
65
66  print "_____\n";
67
68  // Task 5
69  // OLS
70  file2 = "michelson_binary.dat";
71  formula = "Speed ~ -1 + Run";
72  call ols(file2,formula);
73
74  closeall;
75  end;
76
```

Output

```
            Expt
             Run
           Speed
        100.0000

-------------------------------------------------------------------------------------------
Variable        Mean     Std Dev      Variance     Minimum     Maximum     Valid Missing
-------------------------------------------------------------------------------------------

Expt               3       1.421          2.02           1           5       100    0
Run             10.5       5.795         33.59           1          20       100    0
Speed      2.999e+05       79.01          6243   2.996e+05   3.001e+05       100    0

-------------------------------------------------------------------------------

Variance for Speed in Expt 1 is       10459.0000
Variance for Speed in Expt 5 is        2792.7500
-------------------------------------------------------------------------------

t test for both sub samples, where s1 = s2 = 7500 :          2.8299
t test for both sub samples, where s1 != s2 using results from procedure:          3.0108
Since abs(t) > 2.05, for both approaches, it follows means are different
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------

Valid cases:                  100      Dependent variable:               Speed
Missing cases:                  0      Deletion method:                   None
Total SS:        8991146796600.010      Degrees of freedom:                  99
R-squared:                   0.768      Rbar-squared:                     0.768
Residual SS:     2083366316960.802      Std error of est:            145065.862
F(1,99):                   328.253      Probability of F:                 0.000

                        Standard                Prob   Standardized  Cor with
Variable     Estimate      Error      t-value     >|t|     Estimate    Dep Var
-------------------------------------------------------------------------------
Run           21940.3     1210.99     18.1177     0.000     0.87652     0.87652
» |
```