Interpreting SHAZAM Output *to accompany*

Hill, Griffiths, Judge, *Undergraduate Econometrics*, Second Edition

Chapters 3 to 6 discuss the simple linear regression model with application to the estimation and analysis of the food expenditure function. The data set is listed in Table 3.1 on page 50. The SHAZAM commands below load the data set and assign variable names with the **READ** command. The **OLS** command is used for least squares estimation of the linear regression model.

```
SAMPLE 1 40
READ (TAB3-1.txt) FOOD INCOME / SKIPLINES=1
OLS FOOD INCOME / LIST PCOV
* Calculate a confidence interval for the slope parameter
CONFID INCOME
* Test the null hypothesis H0: B2 = 0.10
TEST INCOME=0.10
STOP
```

The first part of the SHAZAM output follows.

```
|_SAMPLE 1 40
|_READ (TAB3-1.txt) FOOD INCOME / SKIPLINES=1
UNIT 88 IS NOW ASSIGNED TO: TAB3-1.txt
   2 VARIABLES AND       40 OBSERVATIONS STARTING AT OBS       1

|_OLS FOOD INCOME / LIST PCOV
 OLS ESTIMATION
       40 OBSERVATIONS     DEPENDENT VARIABLE= FOOD
...NOTE..SAMPLE RANGE SET TO:      1,     40

 R-SQUARE =   0.3171     R-SQUARE ADJUSTED =   0.2991
VARIANCE OF THE ESTIMATE-SIGMA**2 =   1429.2
STANDARD ERROR OF THE ESTIMATE-SIGMA =   37.805
SUM OF SQUARED ERRORS-SSE=   54311.
MEAN OF DEPENDENT VARIABLE =   130.31
LOG OF THE LIKELIHOOD FUNCTION = -201.030

VARIABLE   ESTIMATED  STANDARD   T-RATIO        PARTIAL STANDARDIZED ELASTICITY
  NAME    COEFFICIENT   ERROR      38 DF   P-VALUE CORR. COEFFICIENT  AT MEANS
INCOME    0.12829     0.3054E-01   4.201     0.000 0.563     0.5631     0.6872
CONSTANT   40.768      22.14       1.841     0.073 0.286     0.0000     0.3128
```

The standard error for the slope coefficient is reported as `0.3054E-01`. This means $(0.3054)(10^{-1}) = 0.03054$. It should also be noted that the SHAZAM output typically reports more decimal places of accuracy than are needed for report presentation.

The least squares estimation results listed on the SHAZAM output are interpreted as follows.

VARIABLE NAME	ESTIMATED COEFFICIENT	STANDARD ERROR	T-RATIO 38 DF P-VALUE	. . .	ELASTICITY AT MEANS
INCOME	b_2	$se(b_2)$	$H_0: \beta_2 = 0$ $t = b_2 / se(b_2)$	. . .	$b_2(\bar{x} / \bar{y})$
CONSTANT	b_1	$se(b_1)$	$H_0: \beta_1 = 0$ $t = b_1 / se(b_1)$	. . .	

The least squares estimates match the numerical results reported in Equation (3.3.9), page 55. (Note that the SHAZAM output reports the intercept estimate as the final parameter estimate in the row with the label CONSTANT). The income elasticity evaluated at the sample means is described in Equation (3.3.14), p. 57. The standard errors are given on page 82.

The test of significance reported in the T-RATIO column is discussed on pages 105-106. SHAZAM reports p-values for a 2-sided test. The calculation of the p-value for this example is illustrated in Figure 5.6, p. 106. The SHAZAM ouput reports 3 decimal places of accuracy for the p-value. Therefore, a value of 0.000 means the p-value is less than 0.0005. This leads to the decision that the null hypothesis of a zero coefficient can be rejected at any reasonable significance level.

The SHAZAM least squares (OLS) estimation output reports the following statistics.

		Textbook reference
R-SQUARE	R^2	page 124
VARIANCE OF THE ESTIMATE-SIGMA**2	$\hat{\sigma}^2$	page 82
STANDARD ERROR OF THE ESTIMATE-SIGMA	$\hat{\sigma}$	
SUM OF SQUARED ERRORS-SSE	$SSE = \sum \hat{e}_t^2$	page 124
MEAN OF DEPENDENT VARIABLE	$\bar{y}$	page 55

The **PCOV** option on the **OLS** command reports the estimated variances and covariances of the least squares estimators as follows.

```
VARIANCE-COVARIANCE MATRIX OF COEFFICIENTS
INCOME    0.93265E-03
CONSTANT -0.65099       490.12
             INCOME       CONSTANT
```

The numerical results correspond to the presentation in page 82 of the text and are interpreted as follows.

VARIANCE-COVARIANCE MATRIX OF COEFFICIENTS

INCOME $\widehat{\text{var}}(b_2)$

CONSTANT $\widehat{\text{cov}}(b_1, b_2)$ $\widehat{\text{var}}(b_1)$

INCOME CONSTANT

The **LIST** option on the **OLS** command gives the following additional output that can be compared with Table 4.2, page 82.

t	y_t	$\hat{y}_t = b_1 + b_2 x_t$	$\hat{e}_t = y_t - \hat{y}_t$	residual plot

```
    OBS.    OBSERVED      PREDICTED    CALCULATED
     NO.     VALUE         VALUE        RESIDUAL
      1     52.250        73.905       -21.655                 *  I
      2     58.320        84.783       -26.463                 *  I
      3     81.790        95.290       -13.500                  * I
      4     119.90        100.74        19.158                    I  *
      5     125.80        102.72        23.082                    I  *
      6     100.46        103.33       -2.8739                    *
      7     121.51        104.46        17.047                    I *
      8     100.08        107.40       -7.3207                   *I
      9     127.75        110.47        17.283                    I *
     10     104.94        111.16       -6.2195                   *I

           more output for observations 11 to 40.

 RESIDUAL SUM = -0.11937E-11  RESIDUAL VARIANCE =   1429.2
```

The final calculations listed above are obtained as:

RESIDUAL SUM = $\sum \hat{e}_t$ RESIDUAL VARIANCE = $\hat{\sigma}^2$

The OLS estimated residuals should sum to zero. The output reports the number `-0.11937E-11`. This shows that computer calculations are subject to rounding error that can vary from machine to machine.

The **CONFID** command is used to obtain confidence interval estimates using the results from the previous estimation command. The variable names specified on the **CONFID** command represent coefficients. SHAZAM output for the food expenditure example follows.

```
|_* Calculate a confidence interval for the slope parameter
|_CONFID INCOME
USING 95% AND 90% CONFIDENCE INTERVALS
CONFIDENCE INTERVALS BASED ON T-DISTRIBUTION WITH  38 D.F.
     - T CRITICAL VALUES =   2.021 AND   1.684
NAME   LOWER 2.5%   LOWER 5%   COEFFICIENT   UPPER 5%   UPPER 2.5%   STD. ERROR
INCOME    0.6657E-01  0.7686E-01  0.12829      0.1797      0.1900       0.031
```

The above output shows a 95% interval estimate for the slope coefficient as reported in Equation (R5.2) on page 98. When the degrees of freedom exceeds 30 the SHAZAM program does not calculate t-distribution critical values for every case. In this example, the degrees of freedom is 38. For a tail area probability of 5%, SHAZAM uses a critical value of 2.021 that corresponds to 40 degrees of freedom. This approximation should not alter the conclusions. Take careful note of the interpretation of an interval estimate discussed on page 98.

The **TEST** command can be used for hypothesis testing. The variable names specified on the **TEST** command represent coefficients from the previous estimation results. The SHAZAM output below shows the results from the hypothesis test discussed on pages 102-105.

```
|_* Test the null hypothesis H0: B2 = 0.10
|_TEST INCOME=0.10
TEST VALUE =  0.28289E-01 STD. ERROR OF TEST VALUE  0.30539E-01
T STATISTIC =  0.92630295     WITH   38 D.F.    P-VALUE= 0.36013
```

The above calculations are obtained as:

TEST VALUE = $b_2 - 0.10$ STD. ERROR OF TEST VALUE $se(b_2)$

T STATISTIC = $t = \dfrac{b_2 - 0.10}{se(b_2)}$

SHAZAM reports p-values for a 2-sided test. The p-value for this hypothesis test is illustrated in Figure 5.5, page 105.