Reading Output of Stata regress Command

TOPIC: **Interpreting Output of Stata regress Command**

DATA: **auto1.dta** (a Stata-format data file)

MODEL: $\text{price}_i = \beta_1 + \beta_2\,\text{weight}_i + u_i \quad (i = 1, ..., N)$

```
. regress price weight

  Source |       SS       df       MS                  Number of obs =      74
---------+------------------------------               F(  1,    72) =   29.42
   Model |  184233937     1   184233937                Prob > F      =  0.0000
Residual |  450831459    72  6261548.04                R-squared     =  0.2901
---------+------------------------------               Adj R-squared =  0.2802
   Total |  635065396    73  8699525.97                Root MSE      =  2502.3

------------------------------------------------------------------------------
   price |      Coef.   Std. Err.       t     P>|t|       [95% Conf. Interval]
---------+--------------------------------------------------------------------
  weight |   2.044063   .3768341      5.424   0.000       1.292858    2.795268
   _cons |  -6.707353    1174.43     -0.006   0.995       -2347.89    2334.475
------------------------------------------------------------------------------
```

Source	SS	df	MS = **SS/df**
Model	184233937 = **ESS**	1 = **k–1**	184233937 = **ESS/(k-1)**
Residual	450831459 = **RSS**	72 = **N–k**	6261548.04 = **RSS/(N-k)** = $\hat{\sigma}^2$
Total	635065396 = **TSS**	73 = **N–1**	8699525.97 = **TSS/(N-1)** = S_Y^2

Number of obs = 74 = **N**
F(1, 72) = 29.42 = **F-statistic for test of** H_0: $\beta_2 = 0$ **against** H_1: $\beta_2 \neq 0$
Prob > F = 0.0000 = **p-value for F-statistic**
R-squared = 0.2901 = R^2
Adj R-squared = 0.2802 = $\bar{R}^2$
Root MSE = 2502.3 = $\hat{\sigma}$

price	Coef. = $\hat{\beta}_j$	Std. Err. = $\widehat{se}(\hat{\beta}_j) = \sqrt{\widehat{Var}(\hat{\beta}_j)}$
weight	2.044063 = $\hat{\beta}_2$	.3768341 = $\widehat{se}(\hat{\beta}_2) = \sqrt{\widehat{Var}(\hat{\beta}_2)}$
_cons	-6.707353 = $\hat{\beta}_1$	1174.43 = $\widehat{se}(\hat{\beta}_1) = \sqrt{\widehat{Var}(\hat{\beta}_1)}$

The printed t-statistics are those for performing **two-tail t-tests** of the null hypothesis H_0: $\beta_j = 0$ against the alternative hypothesis H_1: $\beta_j \neq 0$.

- The **sample value of each t-statistic** is the **t-ratio**:

$$t_j = \frac{\hat{\beta}_j}{\hat{se}(\hat{\beta}_j)} = \text{t-ratio for } \hat{\beta}_j \quad (j = 1,2).$$

- The **null distribution of t_j** under H_0: $\beta_j = 0$ is the **t[N-2] distribution**.

- The column labelled "P>|t|" contains the **two-tailed p-values for the t-ratios t_j.**

```
------------------------------
```

t = $t_j = \hat{\beta}_j / \hat{se}(\hat{\beta}_j)$	P>|t| = $\Pr(\lvert t\rvert > \lvert t_j\rvert) = 2\Pr(t > t_j)$
5.424 = $t_2 = \hat{\beta}_2 / \hat{se}(\hat{\beta}_2)$	0.000 = $\Pr(\lvert t\rvert > \lvert t_2\rvert) = 2\Pr(t > t_2)$
-0.006 = $t_1 = \hat{\beta}_1 / \hat{se}(\hat{\beta}_1)$	0.995 = $\Pr(\lvert t\rvert > \lvert t_1\rvert) = 2\Pr(t > t_1)$

The printed confidence intervals are the **two-sided 95 percent confidence intervals for each regression coefficient β_j** (j = 1,2).

- In general, the **two-sided 100(1-α) percent confidence interval for regression coefficient β_j** is:

$$\left[\hat{\beta}_j - t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_j),\ \hat{\beta}_j + t_{\alpha/2}[N-2]\hat{se}(\hat{\beta}_j)\right].$$

- For the **two-sided 95 percent confidence intervals**, 1-α = 0.95, α = 0.05, and α/2 = 0.025.

[95% Conf. Interval]	$\left[\hat{\beta}_j - t_{0.025}[72]\hat{se}(\hat{\beta}_j),\ \hat{\beta}_j + t_{0.025}[72]\hat{se}(\hat{\beta}_j)\right]$
1.292858 = $\hat{\beta}_2 - t_{0.025}[72]\hat{se}(\hat{\beta}_2)$	2.795268 = $\hat{\beta}_2 + t_{0.025}[72]\hat{se}(\hat{\beta}_2)$
-2347.89 = $\hat{\beta}_1 - t_{0.025}[72]\hat{se}(\hat{\beta}_1)$	2334.475 = $\hat{\beta}_1 + t_{0.025}[72]\hat{se}(\hat{\beta}_1)$

Note: $t_{0.025}[72] = 1.9935$