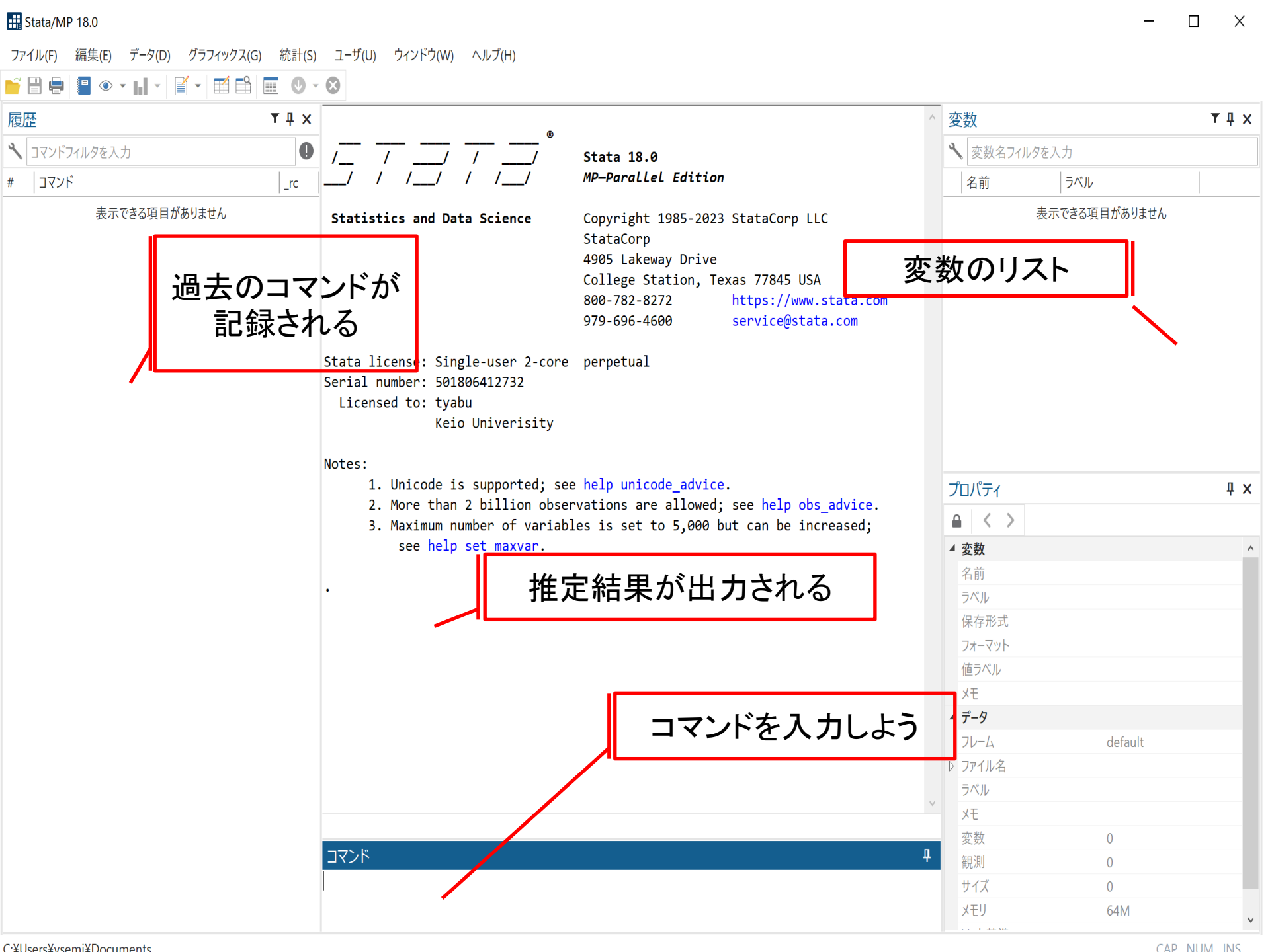


StataのABC



過去のコマンドが
記録される

推定結果が出力される

コマンドを入力しよう

変数のリスト

OLS推定

- OLSのコマンド

reg Y X1 X2 X3

--- 定数項を入れないならnoconstantとする

reg Y X1 X2 X3, noconstant

--- ロバスト標準誤差を使うなら

reg Y X1 X2 X3, r

ロバスト標準誤差のコマンド

- インターフェースを使って同じことができる

統計タブ→線形モデル他→線形回帰

例： 賃料の決定要因

. reg rent distance

Source	SS	df	MS	Number of obs	=	724
Model	142.186468	1	142.186468	F(1, 722)	=	29.34
Residual	3498.69484	722	4.84583772	Prob > F	=	0.0000
				R-squared	=	0.0391
				Adj R-squared	=	0.0377
Total	3640.88131	723	5.0357971	Root MSE	=	2.2013

rent	Coefficient	Std. err.	t	P> t	[95% conf. interval]	
distance	.0929731	.0171638	5.42	0.000	.0592763	.12667
_cons	6.032085	.2025719	29.78	0.000	5.634384	6.429785

- Coefficient: OLS推定した値(_consは定数項、distanceはX)
- Std. err.: 標準誤差
- tは、帰無仮説: パラメータが0として計算したt値
- P>|t|は、p値(0.01以下なら1%有意、0.05以下なら5%有意、0.1以下なら10%有意、0.1より大きいなら有意でない)
- R-squared: 決定係数、Adj R-squared: 自由度調整済み
- F(1, 722): 全ての係数が0としたF値

2SLS推定

- 1st Stage

`reg X Z W1 W2, r`

--- 内生変数Xと操作変数Zが相関しているか

- 2nd Stage

`ivreg Y (X=Z) W1 W2, r`

--- `ivreg Y (X=Z) W1 W2, first r`

(firstと入れると1st Stepの結果も表示してくれる)

(例) 出産と労働時間との関係

・変数の定義

morekids: 子供が2人より多い母親なら1、子供が2人だけなら0

boy1st: 最初の子供が男なら1、 boy2nd: 二番目が男なら1

samesex: 最初の2人が同性なら1

agem1: 母親の年齢

black: 母親が黒人なら1、 hispan: 母親がヒスパニックなら1

othrace: 母親が黒人、ヒスパニック、白人ではない(アジア人など)

weeksm1: 母親の労働時間(1979年に何週働いたか)

Variable	Obs	Mean	Std. Dev.	Min	Max
morekids	254,654	.3805634	.4855263	0	1
boy1st	254,654	.5143607	.4997947	0	1
boy2nd	254,654	.5125504	.4998434	0	1
samesex	254,654	.5055683	.49997	0	1
agem1	254,654	30.39327	3.386447	21	35
black	254,654	.0516623	.2213447	0	1
hispan	254,654	.0742066	.2621073	0	1
othrace	254,654	.0563431	.2305836	0	1
weeksm1	254,654	19.01833	21.86728	0	52

ivreg weeksm1 (morekids=samesex) agem1 black hispan othrace, first r
(firstと入れると1st Stepの結果も表示してくれる)

First-stage regressions						
Source	SS	df	MS	Number of obs = 254654		
Model	1451.65094	5	290.330189	F(5,254648) = 1262.09		
Residual	58579.1858254648	.230039843		Prob > F = 0.0000		
Total	60030.8368254653	.235735832		R-squared = 0.0242		
				Adj R-squared = 0.0242		
				Root MSE = .47962		

morekids	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
agem1	.0153898	.0002814	54.68	0.000	.0148382	.0159415
black	.1005237	.0043058	23.35	0.000	.0920844	.1089629
hispan	.1512258	.0040119	37.69	0.000	.1433625	.1590892
othrace	.0275069	.0045537	6.04	0.000	.0185817	.0364321
samesex	.0680081	.001901	35.77	0.000	.0642821	.0717341
_cons	-.1395319	.0086985	-16.04	0.000	-.1565808	-.122483

- 1st Stepでは、操作変数だけでなく、すべての外生変数が含まれている
- samesexは正で有意である
- 女性の年齢が高いほど子供の数も多い
- ヒスパニック、黒人、アジア人、白人の順で子供が多い

Instrumental variables (2SLS) regression

Number of obs = 254654

F(5,254648) = 1390.96

Prob > F = 0.0000

R-squared = 0.0437

Root MSE = 21.385

weeksm1	Coef.	Robust Std. Err.	t	P> t	[95% Conf. Interval]	
morekids	-5.821051	1.246401	-4.67	0.000	-8.263963	-3.378139
agem1	.8315975	.0226408	36.73	0.000	.7872221	.875973
black	11.62327	.231798	50.14	0.000	11.16896	12.07759
hispan	.4041802	.2607993	1.55	0.121	-.1069795	.9153399
othrace	2.130962	.2109882	10.10	0.000	1.717431	2.544493
_cons	-4.791894	.3897914	-12.29	0.000	-5.555874	-4.027913

Instrumented: morekids

Instruments: agem1 black hispan othrace samesex

--- morekidsの係数は-5.82である

--- 女性の年齢があがると労働時間は増える

--- 黒人、アジア人、ヒスパニック、白人の順で労働時間が多い