Empirical Industrial Organization & Consumer Choice

1c Standard Errors and Hypothesis Tests

Summer 2018

Sebastian Kranz, Uni Ulm

March 29, 2018

An estimator's variance and standard error

- Recall that an estimator $\hat{\beta}$ is a (vector-valued) random variable whose value depends on the realization of the disturbances ε.
- If we know all the parameters of the model, we can compute the variance-covariance matrix of $\hat{\beta}$.
- The variance-covariance matrix of the OLS estimator in the linear model is given by

$$Var(\hat{\beta}) = \sigma^2 (X'X)^{-1}$$

  where σ is the standard deviation of the disturbance ε_t.
- The standard deviation of $\hat{\beta}_i$ is called standard error $se(\hat{\beta}_i)$. It is given by the root of the i'th diagonal element of $\hat{\beta}$.

Estimating an estimator's covariance matrix

- In empirical studies, one is interested in standard errors in order to know how precise an estimate $\hat{\beta}_i$ is.
- By analyzing the residuals $\hat{\varepsilon}$ of the regression, we can get the following unbiased estimate of the variance of the disturbance ε_t:

$$\hat{\sigma}^2 = \frac{1}{T-(K+1)}\sum \hat{\varepsilon}_t^2$$

- An estimator of the variance-covariance matrix of the OLS estimator is then given by

$$\hat{Var}(\hat{\beta}) = \hat{\sigma}^2(X'X)^{-1}$$

  - If A2 is violated, i.e. we have heteroscedastic or autocorrelated disturbances, we need a different, *robust* estimator for $Var(\hat{\beta})$.
  - For an IV estimator $\hat{\beta}^{IV}$ an estimate of the variance-covariance matrix is given by

$$\hat{Var}(\hat{\beta}^{IV}) = \hat{\sigma}^2(X'P_Z X)^{-1}$$

with$P_Z = Z(Z'Z)^{-1}Z'$

Optional: Exploring standard errors

- Extend our R function ols for a parameter compute.se that is by default FALSE. If compute.se is TRUE, the function shall also return the estimated standard errors of the OLS estimates.
- Compare our estimated standard errors with the standard errors that are shown when you estimate and summarize the model with the R functions lm and summary.
- Optional: Show in a simulation study that the mean of the estimated standard errors of the estimated coefficients is roughly equal to the standard deviation of the estimated coefficients in the simulation.

Hypothesis tests: Null hypothesis

- A hypothesis test consists of a **null hypothesis** (H_0) and a corresponding **alternative hypothesis** H_1 about some features of a data generating process. Examples for hypotheses in a linear regression model:
  - H_0: $\beta_1 = 0$, $H_1 : \beta_1 \neq 0$ (a two-sided test)
  - H_0: $\beta_1 \leq 5$, $H_1 : \beta_1 > 5$ (a one-sided test)
  - H_0: The explanatory variable x_k is exogenous
    H_1 : x_k is endogenous
  - H_0: The disturbance ε is not auto-correlated
    H_1: ε is auto-correlated

Hypothesis tests: test statistic

- Hypothesis tests are based on a **test statistic**
  - A test statistic TS is a random variable for which we know the distribution if the null hypothesis H_0 holds true and additionally assumptions are satisfied by the true model (like A1-A3 in the linear regression model).
  - We can compute a realization of the test statistic *ts* from the observed data
- Remarks:
  - If H_0 is an inequality (a one-sided test), like $\beta_1 \leq 0$, the relevant distribution of a test statistic is given by the case that the true parameter satisfies it with equality: $\beta_1 = 0$.
  - It is important to look at a detailed description of a test to learn the assumptions on the data generating process under which the test is valid

Hypothesis tests: p-value

- The **p-value** of a test given a realization ts of the test statistic is the probability to get such or a "more extreme realization" of TS if H_0 (and all other made assumptions) hold true.
- The exact formula of the p-value depends on whether we have a one-tailed or two-tailed test:

$$p = \Pr(TS \geq ts|H_0) \text{ one-tailed test}$$

$$p = 2 * \min\{\Pr(TS \geq ts|H_0), \Pr(TS \leq ts|H_0)\} \text{two-tailed test}$$

- We can **reject** a null hypothesis H_0 at a **significance level** α (e.g. $\alpha = 5\%$) if the p-value is below α.

t-test for a regression coefficient

- Consider a linear regression model that satisfies (A1)-(A3)
- Consider the null hypothesis that the true value of $\hat{\beta}_i$ is 0

$$H_0 : \beta_i = 0$$

- The t-value for an estimated coefficient $\hat{\beta}_i$ is given by

$$t_i = \frac{\hat{\beta}_i}{\hat{se}(\hat{\beta}_i)}$$

- It is a long known result in statistics that if $\beta_i = 0$, then the t-value of the OLS estimator $\hat{\beta}_i$ is distributed according to a t-distribution with $T - K$ degrees of freedom.

Exploring t-tests

- For some value of $T-K$ use R to plot the density of a t-distribution with $T-K$ degrees of freedom.
- The t-test for $H_0 : \beta_i = 0$ is a two-tailed test. Draw vertical lines at the two possible realizations of t_i, for which that t-test has a p-value of 5%.
- Write a function that simulates and estimates via OLS a linear regression model $y = \beta_0 + \beta_1 x_1 + \varepsilon$. The function shall return for each coefficient β_1, the estimated coefficient $\hat{\beta}_i$, the estimated standard error $\hat{se}(\hat{\beta}_i)$, the t-value t_i and the corresponding p-value.
- Check that your t-values and p-values are the same than those you get when estimating the model with lm and analyzing the result with summary.

Exploring t-tests

- Run a simulation study, in which you 1000 times repeat the simulation and estimation. Which distribution for the p-values would you expect when $\beta_1 = 0$? Compare the resulting empirical distribution for the p-values for β_1 if for the cases that $\beta_1 = 0$ and $\beta_1 = 1$. How often do you reject (not reject) H_0 at significance level $\alpha = 5\%$ when $\beta_1 = 0$ and when $\beta_1 = 1$?
- Does the distribution of the p-values depend on the number of observations T in each simulated data set?
- Assume x_1 is endogenous by being positively correlated with ε. Does the t-test for $\beta_1 = 0$ still work correctly?

Some thoughts about hypothesis testing

- Assume the null hypothesis $H_0 : \beta_i = 0$ is not rejected at a significance level $\alpha = 0.1\%$. Is this strong evidence that H_0 is true?
- Do you think there is a statistical method that would allow us to get strong evidence that a null hypothesis of the form $H_0 : \beta_i = 0$ is true?
- Assume null hypothesis $H_0 : \beta_i = 0$ is rejected at a significance level $\alpha = 0.1\%$ ($p \leq 0.001$). Is this strong evidence that H_0 is false?
- Assume $H_0 : \beta_i = 0$ is rejected at a significance level $\alpha = 5\%$ ($p \leq 0.05$). Does it mean that the probability that H_0 is true is equal to or smaller than 5%?
    - http://xkcd.com/1132/
- If you just run a lot of different regressions, you are likely to find some significant results by chance:
    - http://xkcd.com/882/

F-test for joint significance

- Run in R an OLS regression with two or more explanatory variables

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_K x_K + \varepsilon$$

- A `summary` of the regression results shows in R the results of a F-test for joint significance of all regressors, e.g.
  - `F-statistic: 104 on 1 and 94 DF, p-value:` $<$ `2.2e-16`
- The Null Hypothesis of the F-test is

$$H_0 : \beta_1 = \beta_2 = \ldots = \beta_K = 0$$

  i.e. H_0 is that no explanatory variable has an influence on y.
- For details of how the F-statistic is computed, see e.g. http://en.wikipedia.org/wiki/F-test

Diagnostic tests for IV estimation

- Run an instrumental variable estimation with R and show a `summary` of the results with the option `diagnostic=TRUE`
- You see results of 3 diagnostic tests:
  - Weak instruments
  - Wu-Hausman (endogeneity of regressors)
  - Sargan (endogeniety of instruments)
- Unfortunately, currently the R-help for summary.ivreg provides almost no information what the tests do and how we should interpret these results. We will very briefly give an overview over these tests.
  - Side remark: even though R has many benefits, Stata is currently much stronger in documentation and incorporation of econometric tests and options for estimation.

Testing for weak instruments

- Consider a simple regression model of a demand function

$$q_t = \beta_0 + \beta_1 p_t + \beta_2 s_2 + \varepsilon_t$$

  with endogenous prices p and an exogenous, explanatory variables s_2 and excluded instruments z_1,z_2.
- The weak instruments problem means that if the instruments are only weakly correlated with p the IV estimator can become considerably biassed (and imprecise) for small sample size T.
- The test for weak instruments shown in R is an F-test on the joint signficance of all explanatory variables in the first stage regression of the two stage least squares procedure:

$$p = \beta_2 s_2 + \gamma_1 z_1 + \gamma_2 z_2 + \eta$$

- Rule-of-thumb: Staiger and Stock (1997) suggested declaring instruments to be weak if the first-stage F-statistic is smaller than 10, Stock and Yogo (2005) provide much more details.

Wu-Hausman test for endogenous regressors

- Consider a linear regression model of a demand function

$$q_t = \beta_0 + \beta_1 p_t + \varepsilon_t$$

  for which we don't know if prices p are endogenous or exogenous.
- If we have valid instruments z for a possibly endogenous variable p, the *Wu-Hausman test* allows to test whether p is indeed endogenous.
- The Null hypothesis of the Wu-Hausman test is that p is exogenous
  - i.e. low p-values of the Wu-Hausman test suggest an endogenous variable

Wu-Hausman test for endogenous regressors

- Consider the regression

$$y = X^{exo}\alpha + X^{endo}\beta + \varepsilon$$

- X^{endo} is the matrix or vector of variables that are **possibly** endogenous
- Null Hypothesis: All variables in X^{endo} are exogenous.
- Let Z be a matrix of valid instruments (including X^{exo})
- Let $\hat{X}^{endo}$ be the fitted values from the first stage regression of 2SLS, in which all variables in X^{endo} are regressed on Z.
- The Wu-Hausman test is equal to an F-test on the null hypothesis that $\gamma = 0$ in the regression

$$y = X^{exo}\alpha + X^{endo}\beta + \hat{X}^{endo}\gamma + \varepsilon$$

- A more detailed description of the Wu-Hausman test, is e.g. given in Green's textbook "Econometrics Analysis".

Sargan test for endogenous instruments

- The Sargan test is a test with the Null hypothesis that all instruments are exogenous.
- The Sargan test can only be applied if we have at least one more excluded instrument than endogenous variable.
- If the Sargan test is rejected (low $p - value$), it suggests that at least one instrument is endogenous.
- But: If the Sargan test is not rejected we do **not** have strong proof that all instruments are indeed exogenous, e.g. the Sargan test may well fail to detect if all instruments are endogenous.
  - This means not being rejected by the Sargan test can be interpreted as a neccessary condition for exogenous instruments but not a sufficient one. Most important remains the economic reasoning behind the selection of the instruments.
- Exercise: Explore the Sargan test in simulation studies.