Measuring Uncertainty in Statistical Estimates

Wednesday, October 31, 2012 11:57 AM

- Statistical models allow us to estimate the relationship between variables of interest
- We know, intuitively, that these estimates are imperfect
- A model should try to minimize these imperfections…
- ...but, if ultimately irresolvable, to measure the extent to which imperfections might cause variability in the results
- What are sources of error in estimation?

1) Sampling variation - error that accrues to results because even properly implemented sampling procedures do not perfectly mirror the underlying population because of chance in who is selected for the sample.

2) Noise in the data-generating process

$$y = \beta_0 + \beta_1 x + \underline{\underline{\varepsilon}} \qquad \varepsilon \sim \Phi(0, \sigma^2)$$

3) Model misspecification

Parametric Approaches to Measuring Uncertainty

Wednesday, October 31, 2012 12:42 PM

- We can measure the overall degree of error in the DGP (e.g., the standard error of the estimate in an OLS process), but more typically we are interested in the degree of uncertainty in some estimated relationship (e.g., a marginal effect)
- As these effects are usually parameterized, uncertainty in the estimate = uncertainty in the parameter
- Uncertainty in the parameter is itself functionally derived and depends on the quality of the function
  - Accurate specification of the model and VCV

1) Difference of means test: variance estimates

$$\mu_1 - \mu_2 \neq 0 \qquad t = \frac{\bar{x}_1 - \bar{x}_2}{se_{dif}}$$

$$\mu_1 = \mu_2 \qquad se_{dif} = \left[\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right]^{1/2}$$

$$\sigma_1 \qquad \sigma_2$$

$y = x\beta + \varepsilon$

2) OLS: VCV matrix

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \ldots + \hat{\beta}_k x_k$$

$$\underline{\underline{VCV}} = \hat{\sigma}^2 (X'X)^{-1} \qquad \frac{1}{n-k} \cdot u'u$$

ε + constant variance?

linearity/ model specification?

3) Maximum likelihood

Nonparametric Approaches to Estimating Uncertainty

Wednesday, October 31, 2012 1:30 PM

- A non-parametric approach to measuring uncertainty might be advisable in situations where:
  - existing parametric solutions are unadvisable or questionable
  - we cannot solve for the appropriate parametric function
- Example: Mann-Whitney/Wilcoxon Rank Sum test (see Stata 11 manual page for more details)

$X \sim Z$

are the medians of X and Z the same?

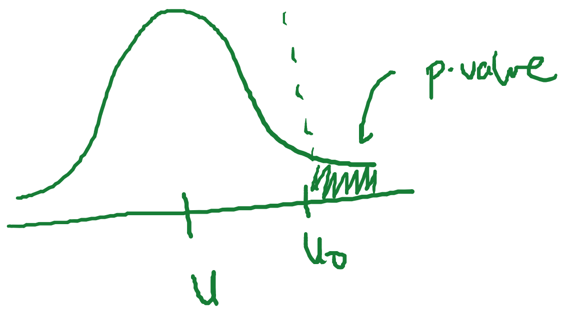
X: $x_1, x_2, x_3, \ldots, x_n$

Z: $z_1, z_2, \ldots, z_m$

$$R(x_1, z_1) = \begin{cases} 1 & \text{if } x_1 > z_1 \\ 0 & \text{if not.} \end{cases}$$

$$U_0 = \sum_j \sum_i R(x_i, z_j)$$

U has a distribution under the null given by combinatorics.


p-value
U_0
U

Bootstrapping

Thursday, November 01, 2012 1:02 PM

- Rather than try to extract information about variability from the estimated error terms $\hat{\epsilon}$, simulate sampling variability
  - Repeatedly simulate a sample of the same size (somehow)
  - Estimate the model on each simulated sample
  - Recover the important parameters from each estimate
  - Plot the distribution of the parameters estimated via simulated sampling
- This is called *bootstrapping*

Parametric Bootstrapping

Wednesday, October 31, 2012 1:44 PM

- Probably best-categorized as a "semi-parametric" approach to estimating uncertainty
- Best applied when attempting to determine the distribution of a complex quantity based on model products
- Example: consider a regression model with an interaction term:

  $$y = \beta_0 + \beta_1 x + \beta_2 z + \beta_p xz + \epsilon$$

  If we estimate this model on a data set, how do we determine the uncertainty of the marginal effect $\frac{\;}{dx}$ when $z \neq 0$?

- One approach: simulate draws from the distribution of $\hat{\beta}$ out of a regression

$$\frac{d\hat{y}}{dx} = \hat{\beta}_1 + \hat{\beta}_p z$$

$$var\left(\frac{d\hat{y}}{dx}\right)$$

$$var(\hat{\beta}_1 + \hat{\beta}_p z)$$

$$var(\hat{\beta}_1) + var(\hat{\beta}_p z) + 2\,cov(\hat{\beta}_1, \hat{\beta}_p z)$$

$$var(\hat{\beta}_1) + z^2\, var(\hat{\beta}_p) + 2z\, cov(\hat{\beta}_1, \hat{\beta}_p)$$

$$f(\hat{\beta}_1, \hat{\beta}_p) \sim \overset{\text{CLT}}{\hat{N}}(\underbrace{\hat{\beta}}_{\text{fitted MLE coefficients}}, \underbrace{\hat{\Sigma}}_{\text{VCV}})$$

① recover OLS/MLE coefficients / VCV
② plug estimated coefs / VCV into asymptotic dist. of $\hat{\beta}$
③ simulate many draws of $\hat{\beta}$ for asymptotic dist.
④ compute $d\hat{y}/dx$ for each draw.

- This is what the Clarify package of King, Tomz, and Wittenberg does
- There is also a parametric approach. But the parametric approach may not be as facile to employ in, e.g., non-linear models.

Nonparametric Bootstrapping

Wednesday, October 31, 2012 5:15 PM

- We could also take an approach that is less-dependent on parametric models
    - Might be necessary if there *is* no parametric model
    - Might be necessary if we have reason to doubt that the parametric formula for uncertainty might be inapplicable to this situation for some reason
    - Might be easier to implement than parametric bootstrapping

- Instead of sampling out of a model quantity, sample directly out of the data set (with replacement)
    - Mimics the sampling process of gathering a data set
    - The sample data set stands in for "the world" in this case

P	?				
S	x_1	x_2	x_3		
b_1	x_1	x_1	x_3	$\rightarrow$	$\hat{\beta}_1$
b_2	x_3	x_2	x_3	$\rightarrow$	$\hat{\beta}_2$
$\vdots$					$\vdots$
b_k	x_2	x_2	x_2	$\rightarrow$	$\hat{\beta}_k$

1) Start with a data set of size *n*

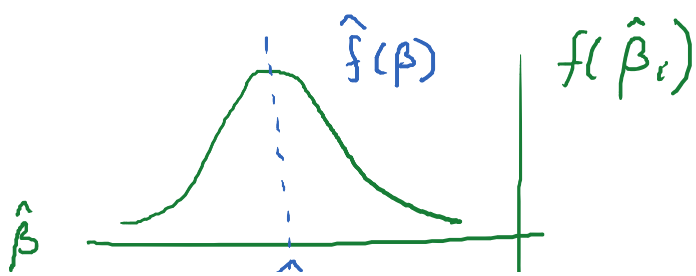
2) Randomly choose *n* draws, with replacement, out of the data set

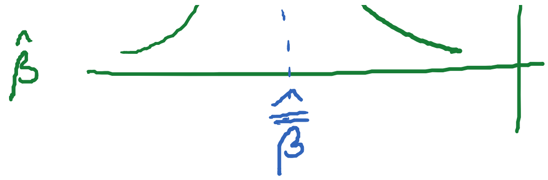
3) Estimate the model on the artificially constructed data set

4) Save relevant model quantities out of the estimated model

5) Repeat steps 2-4 *k* times, where *k* is large

6) Use the distribution of the estimated quantities to compute variances, construct confidence intervals, build density plots, etc.





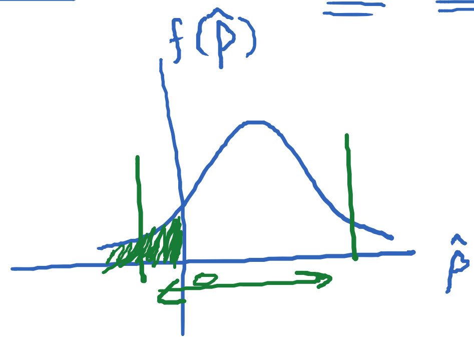
Hypothesis Testing and Confidence Intervals

Wednesday, October 31, 2012 8:21 PM

- We can use the products of bootstrapping in several ways to construct hypothesis tests

1) direct construction of 95% confidence intervals by selecting the 2.5th and 97.5th quantiles of the simulated quantity

percentile CI



2) Using bootstrap estimates of relevant quantities (e.g., of the variance) with a functional approximation (e.g., the t-distribution) to compute confidence intervals

$$\hat{\beta} \pm 1.96\, se(\hat{\beta})$$

estimate this via bootstrapping

$var(\hat{\beta}_i)$

3) Using bootstrapped calculations of a relevant distributional statistic (e.g., the t-statistic) to determine the critical t-values to calculate a confidence interval

bootstrapping this.

Clustering Standard Errors through Randomization/Permutation Testing

Thursday, November 01, 2012 1:13 PM

- Outstanding problem in political methodology: how do we deal with clustered data?
- Example: experimental data with k subjects in each of r sessions, so that $n = kr$
  - Issue: subject responses may be correlated within session in unpredictable ways
  - Treatment variable is typically constant by session
  - Unmodelled within-session inter-subject correlation may interfere with estimation of (in particular) treatment effect variance (in theory can push the estimated variance up or down, but usually naïve estimates of variance will be too small)

- Clustered robust standard errors (CRSEs) perform poorly when r is small (can actually make the variance deflation problem worse!)

- Kelly Rader suggests applying a randomization test to the data
  - Example DGP: $y_{ij} = \beta_0 + \beta_1 x_{ij} + \beta_2 z_j + \epsilon$, j indexes group and i indexes individual
  - Estimate model on sample data set
  - For a large number of replications, randomly reassign values of *z* but ensure that each member of the group received the same value of *z*
  - Compute the test statistic for β_2 for each data set
  - Compare the value of the test statistic in the original sample to the distribution of simulated test statistics (e.g., determine the p-value using the quantiles of the simulated distribution)