Small-Sample Inference and Bootstrap

Leonid Kogan

MIT, Sloan

15.450, Fall 2010

Outline

Overview

- So far, our inference has been based on asymptotic results: LLN and CLT.
- Asymptotic inference is sometimes difficult to apply, too complicated analytically.
- In small samples, asymptotic inference may be unreliable:
  - Estimators may be consistent but biased.
  - Standard errors may be imprecise, leading to incorrect confidence intervals and statistical test size.
- We can use simulation methods to deal with some of these issues:
  - Bootstrap can be used instead of asymptotic inference to deal with analytically challenging problems.
  - Bootstrap can be used to adjust for bias.
  - Monte Carlo simulation can be used to gain insight into the properties of statistical procedures.

Outline

Example: Autocorrelation

- We want to estimate first-order autocorrelation of a time series x_t (e.g., inflation), $\text{corr}(x_t, x_{t+1})$.
- Estimate by OLS (GMM)

$$x_t = a_0 + \rho_1 x_{t-1} + \varepsilon_t$$

- We know that this estimator is consistent:

$$\text{plim}_{T\to\infty}\, \widehat{\rho}_1 = \rho_1$$

- We want to know if this estimator is biased, i.e., we want to estimate

$$\text{E}(\widehat{\rho}_1) - \rho_1$$

Example: Autocorrelation, Monte Carlo

- Perform a Monte Carlo study to gain insight into the phenomenon.
- Simulate independently N random series of length T.
- Each series follows an AR(1) process with persistence ρ_1 and Gaussian errors:

$$x_t = \rho_1 x_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1)$$

- Compute $\widehat{\rho}_1(n)$, $n = 1, ..., N$ for each simulated sample.
- Estimate the bias:

$$\widehat{\mathrm{E}}(\widehat{\rho}_1) - \rho_1 = \frac{1}{N} \sum_{n=1}^{N} \widehat{\rho}_1(n) - \rho_1$$

Standard error of our simulation-based estimate is

$$\widehat{\sigma} = \sqrt{\frac{1}{N} \sum_{n=1}^{N} \left(\widehat{\rho}_1(n) - \widehat{\mathrm{E}}(\widehat{\rho}_1) \right)^2}$$

Example: Autocorrelation, Monte Carlo

MATLAB ® Code

```
phi = 0.9;                % AR(1) coefficient
T = 100;                  % Sample lengths
N = 100000;               % Number of simulated samples
varss =  1/(1-phi^2); % STD of steady-state distribution
for n=1:N
    x = zeros(T,1);
    x(1) = sqrt(varss)*randn(1,1);          % Draw initial value
    noise = randn(T-1,1);
    for t=2:T
        x(t) = phi*x(t-1) + noise(t-1);
    end
    X = [ones(T-1,1) x(1:T-1)];
    b = (X'*X)\(X'*x(2:T)); rho(n) = b(2); % Run OLS
end
MeanBias = mean(rho) - phi
StdErrorBias = std(rho)/sqrt(N)
```

Example: Autocorrelation, Monte Carlo

- We use 100,000 simulations to estimate the average bias

ρ_1	T	Average Bias
0.9	50	-0.0826 ± 0.0006
0.0	50	-0.0203 ± 0.0009
0.9	100	-0.0402 ± 0.0004
0.0	100	-0.0100 ± 0.0006

- Bias seems increasing in ρ_1, and decreasing with sample size.
- There is an analytical formula for the average bias due to Kendall:

$$E(\widehat{\rho}_1) - \rho_1 \approx -\frac{1 + 3\rho_1}{T}$$

- When explicit formulas are not known, can use bootstrap to estimate the bias.

Example: Predictive Regression

- Consider a predictive regression (e.g., forecasting stock returns using dividend yield)

$$r_{t+1} = \alpha + \beta x_t + u_{t+1}$$
$$x_{t+1} = \theta + \rho x_t + \varepsilon_{t+1}$$
$$(u_t, \varepsilon_t)' \sim \mathcal{N}(0, \Sigma)$$

- Stambaugh bias:

$$\mathrm{E}(\widehat{\beta} - \beta) = \frac{\mathrm{Cov}(u_t, \varepsilon_t)}{\mathrm{Var}(\varepsilon_t)} \mathrm{E}(\widehat{\rho} - \rho) \approx -\frac{1 + 3\rho}{T} \frac{\mathrm{Cov}(u_t, \varepsilon_t)}{\mathrm{Var}(\varepsilon_t)}$$

- In case of dividend yield forecasting stock returns, the bias is positive, and can be substantial compared to the standard error of $\widehat{\beta}$.

Predictive Regression: Monte Carlo

- Predictive regression of monthly S&P 500 excess returns on log dividend yield:

$$r_{t+1} = \alpha + \beta x_t + u_{t+1}$$
$$x_{t+1} = \theta + \rho x_t + \varepsilon_{t+1}$$

- Data: CRSP, From 1/31/1934 to 12/31/2008.
- Parameter estimates:

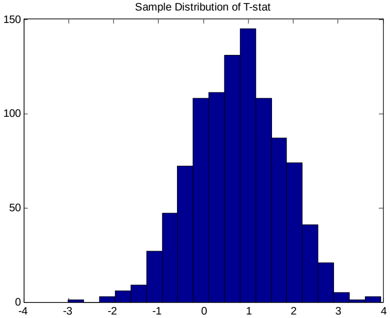
$$\widehat{\beta} = 0.0089, \quad \widehat{\rho} = 0.9936,$$

- S.E.($\widehat{\beta}$) = 0.005.

Predictive Regression: Monte Carlo

- Generate 1,000 samples with parameters equal to empirical estimates. Use 200 periods as burn-in, retain samples of the same length as historical.
- Tabulate $\widehat{\beta}$ and standard errors for each sample. Use Newey-West with 6 lags to compute standard errors.

- Average of $\widehat{\beta}$ is 0.013.
- Average bias in $\widehat{\beta}$ is 0.004.
- Average standard error is 0.005.
- Average t-stat on β is 0.75.



Testing the Mean: Non-Gaussian Errors

- We estimate the mean μ of a distribution by the sample mean. Tests are based on the asymptotic distribution

$$\frac{\widehat{\mu} - \mu}{\widehat{\sigma}/\sqrt{T}} \sim \mathcal{N}(0, 1)$$

- How good is the normal approximation in finite samples if the sample comes from a Non-Gaussian distribution?
- Assume that the sample is generated by a lognormal distribution:

$$x_t = e^{-\frac{1}{2} + \varepsilon_t}, \quad \varepsilon_t \sim \mathcal{N}(0, 1)$$

Lognormal Example: Monte Carlo

- Monte Carlo experiment: N = 100,000, T = 50. Document the distribution of the *t*-statistic

$$\widehat{t} = \frac{\widehat{\mu} - 1}{\widehat{\sigma}/\sqrt{T}}$$

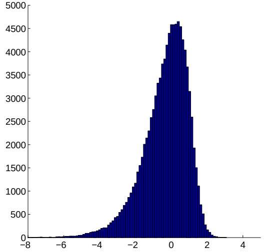
- Asymptotic theory dictates that $\text{Var}(\widehat{t}) = 1$. We estimate

$$\text{Var}(\widehat{t}) = 1.2542^2$$

- Tails of the distribution of $\widehat{t}$ are far from the asymptotic values:

$$\text{Prob}(\widehat{t} > 1.96) \approx 0.0042, \quad \text{Prob}(\widehat{t} < -1.96) \approx 0.1053$$

A histogram of $\widehat{t}$ across 100,000 simulations



Outline

Bootstrap: General Principle

- Bootstrap is a re-sampling method which can be used to evaluate properties of statistical estimators.
- Bootstrap is effectively a Monte Carlo study which uses the empirical distribution as if it were the true distribution.
- Key applications of bootstrap methodology:
    - Evaluate distributional properties of complicated estimators, perform bias adjustment;
    - Improve the precision of asymptotic approximations in small samples (confidence intervals, test rejection regions, etc.)

Bootstrap for IID Observations

- Suppose we are given a sample of IID observations x_t, $t = 1, ..., T$.
- We estimate the sample mean as $\widehat{\mu} = \widehat{\mathrm{E}}(x_t)$. What is the 95% confidence interval for this estimator?
- Asymptotic theory suggests computing the confidence interval based on the Normal approximation

$$\sqrt{T}\,\frac{\widehat{\mathrm{E}}(x_t) - \mu}{\widehat{\sigma}} \sim \mathcal{N}(0,1), \quad \widehat{\sigma}^2 = \frac{\sum_{t=1}^{T}[x_t - \widehat{\mathrm{E}}(x_t)]^2}{T}$$

- Under the empirical distribution, x is equally likely to take one of the values $x_1, x_2, ..., x_T$.

Key Idea of Bootstrap

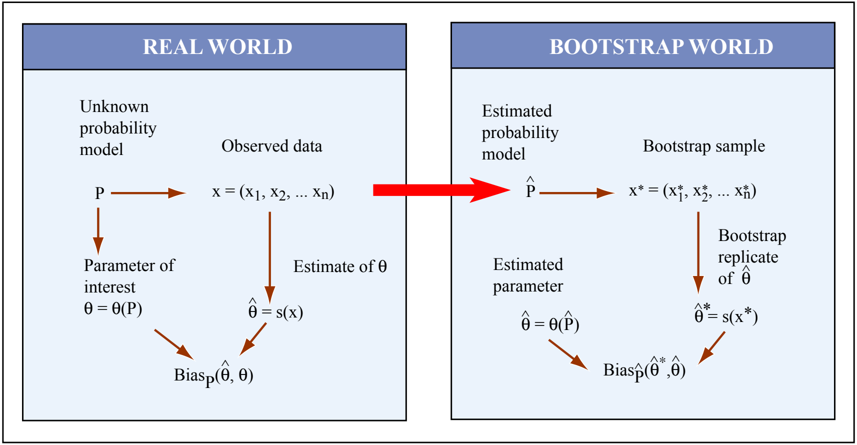


Image by MIT OpenCourseWare.

Source: Efron and Tibshirani, 1994, Figure 10.4.

Bootstrap Confidence Intervals

- Bootstrap confidence interval starts by drawing *R* samples from the empirical distribution.
- For each bootstrapped sample, compute $\widehat{\mu}^{\star}$. "$\star$" denotes statistics computed using bootstrapped samples.
- Compute 2.5% and 97.5% percentiles of the resulting distribution of $\widehat{\mu}^{\star}$:

$$\widehat{\mu}^{\star}_{2.5\%}, \quad \widehat{\mu}^{\star}_{97.5\%}$$

- Approximate the distribution of $\widehat{\mu} - \mu$ with the simulated distribution of $\widehat{\mu}^{\star} - \widehat{\mu}$. Estimate the confidence interval as

$$\left(\widehat{\mu} - (\widehat{\mu}^{\star}_{97.5\%} - \widehat{\mu}),\ \widehat{\mu} - (\widehat{\mu}^{\star}_{2.5\%} - \widehat{\mu})\right)$$

Example: Lognormal Distribution

- Fix a sample of 50 observations from a lognormal distribution $\ln x_t \sim \mathcal{N}(-1/2, 1)$ and compute the estimates

$$\widehat{\mu} = 1.1784, \quad \widehat{\sigma} = 1.5340$$

- Population mean

$$\mu = \mathrm{E}(x_t) = \mathrm{E}(e^{-\frac{1}{2}+\varepsilon_t}) = 1, \quad \varepsilon_t \sim \mathcal{N}(0, 1)$$

- Asymptotic approximation produces a confidence interval

$$(\widehat{\mu} - 1.96\frac{\widehat{\sigma}}{\sqrt{T}}, \ \widehat{\mu} + 1.96\frac{\widehat{\sigma}}{\sqrt{T}}) = (0.7532, \ 1.6036)$$

- Compare this to the bootstrapped distribution.

Lognormal Distribution

Use bootstrap instead of asymptotic inference.

MATLAB® Code

```
R = 10000;
muvec = zeros(R,1);
for r=1:R
    y = x(ceil(T*rand(T,1)));  % Sample with replacement
    muvec(r) = mean(y);
end
muvec = sort(muvec);
% 5 percent confidence interval
LeftEnd = muhat - (muvec(ceil(0.975*R)) - muhat)
RightEnd = muhat - (muvec(floor(0.025*R)) - muhat)
```

Bootstrap estimate of the confidence interval

$$(0.7280,\ 1.5615)$$

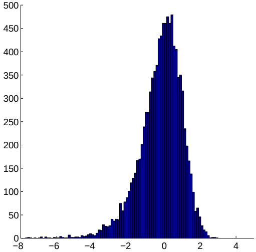
Testing the Mean: Bootstrap

Lognormal Example

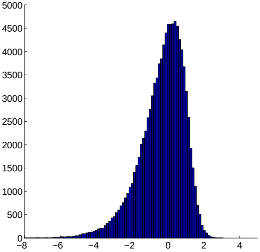
- Consistent with Monte Carlo results: small-sample distribution of t-statistics exhibits left-skewness.
- Variance of the bootstrapped t-statistic is 1.1852^2. Normal approximation: $\text{Var}(\widehat{t}) = 1$. Monte Carlo estimate: $\text{Var}(\widehat{t}) = 1.2542^2$.

A histogram of $\widehat{t}$ statistic

Bootstrap (10,000 samples)



Monte Carlo (100,000 samples)



Boostrap Confidence Intervals

- The basic bootstrap confidence interval is valid, and can be used in situations when asymptotic inference is too difficult to perform.
- Bootstrap confidence interval is as accurate asymptotically as the interval based on the normal approximation.
- For t-statistic, bootstrapped distribution is more accurate than the large-sample normal approximation.
- Many generalizations of basic bootstrap have been developed for wider applicability and better inference quality.

Parametric Bootstrap

- Parametric bootstrap can handle non-IID samples.
- Example: a sample from an AR(1) process: x_t, $t = 1, ..., T$:

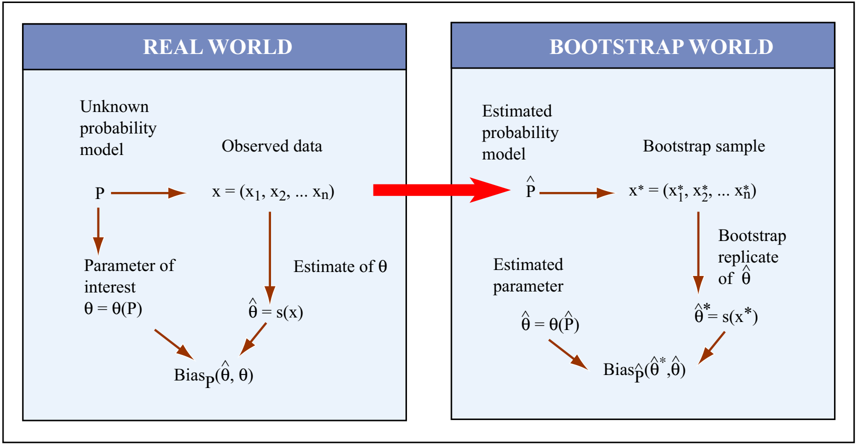
$$x_t = a_0 + a_1 x_{t-1} + \varepsilon_t$$

- Want to estimate a confidence interval for $\widehat{a}_1$.
  1. Estimate the parameters $\widehat{a}_0$, $\widehat{a}_1$ and the residuals $\widehat{\varepsilon}_t$.
  2. Generate R bootstrap samples for x_t.
     - For each sample: generate a long series according to the AR(1) dynamics with $\widehat{a}_0$, $\widehat{a}_1$, drawing shocks with replacement from the sample $\widehat{\varepsilon}_1, ..., \widehat{\varepsilon}_T$;
     - Retain only the last T observations (drop the *burn-in sample*).
  3. Compute the confidence interval as we would with basic nonparametric bootstrap using R samples.

Bootstrap Bias Adjustment

- Want to estimate small-sample bias of a statistic $\widehat{\theta}$:

$$E\left[\widehat{\theta} - \theta_0\right]$$



Source: Efron and Tibshirani, 1994, Figure 10.4.

Bootstrap Bias Adjustment

- Bootstrap provides an intuitive approach:

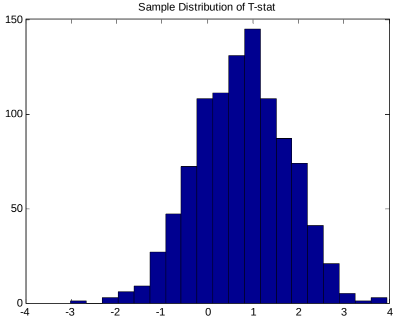
$$\mathrm{E}\left[\widehat{\theta} - \theta_0\right] \approx \mathrm{E}_R\left[\widehat{\theta}^{\star} - \widehat{\theta}\right]$$

  where E_R denotes the average across the R bootstrapped samples.
- Intuition: treat the empirical distribution as exact, compute the average bias across bootstrapped samples.
- Caution: by estimating the bias, we may be adding sampling error. Correct for the bias if it is large compared to the standard error of $\widehat{\theta}$.

Example: Predictive Regression

- Use parametric bootstrap: 1,000 samples, 200 periods as burn-in, retain samples of same length as historical.
- Tabulate $\widehat{\beta}$ and standard errors for each sample. Use Newey-West with 6 lags to compute standard errors.

- Average of $\widehat{\beta}$ is 0.0125.
- Average bias in $\widehat{\beta}$ is 0.0036.
- Average standard error is 0.005.
- Average *t*-stat on β is 0.67.



Discussion

- Asymptotic theory is very convenient when available, but in small samples results may be inaccurate.
- Use Monte Carlo simulations to gain intuition.
- Bootstrap is a powerful tool. Use it when asymptotic theory is unavailable or suspect.
- Bootstrap is *not a silver bullet*:
  - Does not work well if rare events are missing from the empirical sample;
  - Does not account for more subtle biases, e.g., survivorship, or sample selection.
  - Does not cure model misspecification.
- No substitute for common sense!

Readings

- Campbell, Lo, MacKinlay, 1997, Section 7.2, pp. 273-274.
- B. Efron and R.J. Tibshirani, *An Introduction to the Bootstrap*, Sections 4.2-4.3, 10.1-10.2, 12.1-12.5.
- A. C. Davison and D. V. Hinkley, *Bootstrap Methods and Their Application*, Ch. 2. Cambridge University Press, 1997.
- R. Stambaugh, 1999, "Predictive Regressions," *Journal of Financial Economics* 54, 375-421.

15.450 Analytics of Finance
Fall 2010