Example 1 of econometric analysis: the Market Model

Susan Thomas
IGIDR, Bombay

18 November, 2008

Recap: framework for estimating β_i

- We want to estimate the following linear relationship:

$$r_i = r_f + \beta_i(r_M - r_f)$$

Equivalent statement:

$$r_i - r_f = \beta_i(r_M - r_f)$$

- Form of the econometric model:

$$Y_i = \alpha + \beta_i X_i + \epsilon_i$$

- What are the null hypotheses?

$$\begin{aligned}
1.\ \text{If } Y_i &= r_i, X_i = r_M - r_f \\
H_0 : \beta_i &= 1; H_a : \beta \neq 1 \\
H_0 : \alpha &= r_f; H_a : \alpha \neq r_f - \text{ not well posed} \\
2.\ \text{If } Y_i &= r_i - r_f, X_i = r_M - r_f \\
H_0 : \beta_i &= 1; H_a : \beta \neq 1 \\
H_0 : \alpha &= 0; H_a : \alpha \neq 0
\end{aligned}$$

Estimate the market model for a security

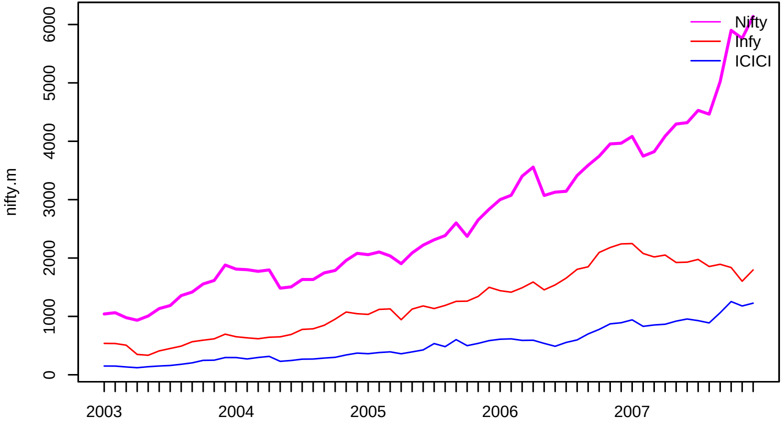
- Time period for analysis: Jan 1 2003 to Dec 31 2007
- Frequency of data: monthly
- Fully diversified portfolio: NSE-50 market index.
- Any stock – it is not important that it is in the stock index or not.
  Our examples: Infosys Technologies, ICICI Bank
- Calculate monthly returns for the stocks and the index:

  $$r_i = 100 * (\log P_1 - \log P_0)$$

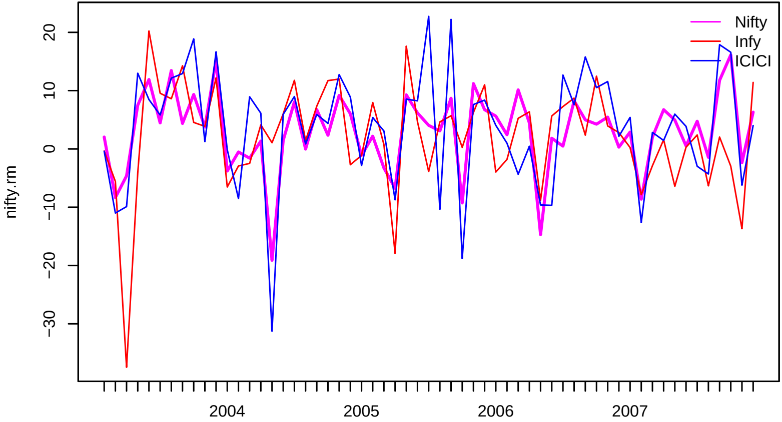
  where P_0 is the price at the start of the month. P_1 is the price at the end of the month.
- r_f is the three-month interest rate set by the RBI.
- Use OLS to estimate beta for Infosys and ICICI Bank assuming the single index market model is the best model for $E(r_i)$.

Regression results for ICICI Bank returns

Recap: Distribution of prices



Recap: Distribution of ∂ returns



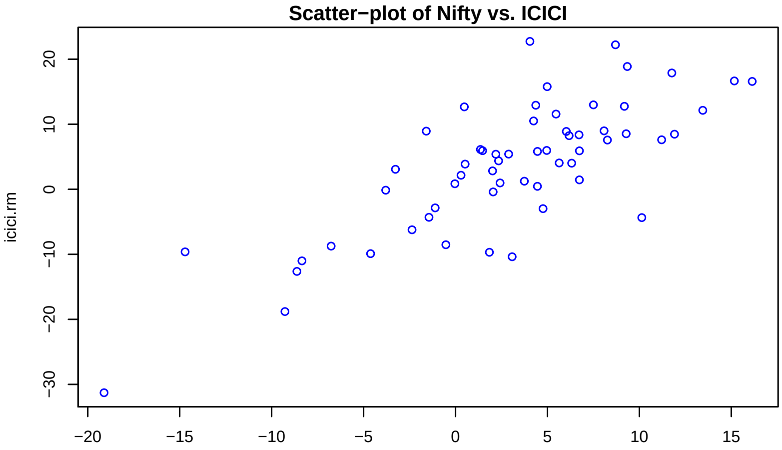
Statistical tests for normality: ICICI Bank

- Summary statistics for monthly data

	ICICI Bank
Mean	3.5629
Std. Dev.	10.1755
Skewness	-0.7185
	(5.16)
Kurtosis	3.975
	(2.377)
N = 60	

- $\chi^2(1)$, 0.05% level of significance = 3.84, 0.01% = 6.63
- $\chi^2(2)$, 0.05% level of significance = 5.99, 0.01% = 9.21
- Significantly different from a normal distribution – will it affect the OLS estimation? How?
- Will MLE be a better estimation tool here? Not if we don't know the exact form of the distribution.

Distribution of Nifty returns vs. ICICI Bank returns



ICICI Bank regression results

- ICICI Bank: $r_{\text{icicibank}} - r_f = \alpha_0 + \beta_{\text{icicibank}}(r_{\text{Nifty}} - r_f) + \epsilon$
- Regression results:

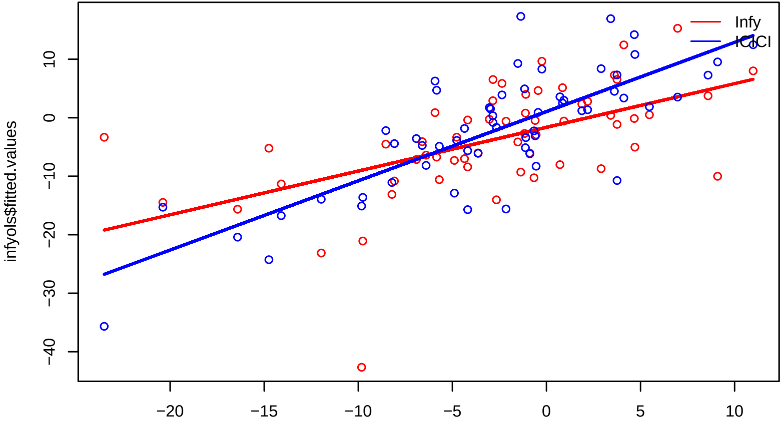
	ICICI Bank		Infosys Tech	
	Estimate	Std. Error	Estimate	Std. Error
(Intercept)	1.0513	0.9030	-1.6413	1.0974
Market returns	1.1837	0.1247	0.7474	0.1516
F-stat(1, 57) =	90.05		24.31	
prob value =	2.5e-13		7.5e-6	
R^2 =	0.6124		0.2989	
Adj R^2 =	0.6056		0.2866	
σ^2_ϵ =	7.757		6.383	

- Model: $E(r_{\text{icicibank}} - r_f) = 1.1837(r_{\text{Nifty}} - r_f)$

Interpreting the regression results

- On average, ICICI Bank price move by 1.2% when the market returns are 1%.
  In comparison: on average, Infosys returns move 0.75% with a 1% move in the market portfolio.
- Market returns captures 61% of the variation in ICICI Bank returns.
  But it captures only 30% of variation in Infosys Technologies.

The regression lines for Infosys Tech and ICICI Bank



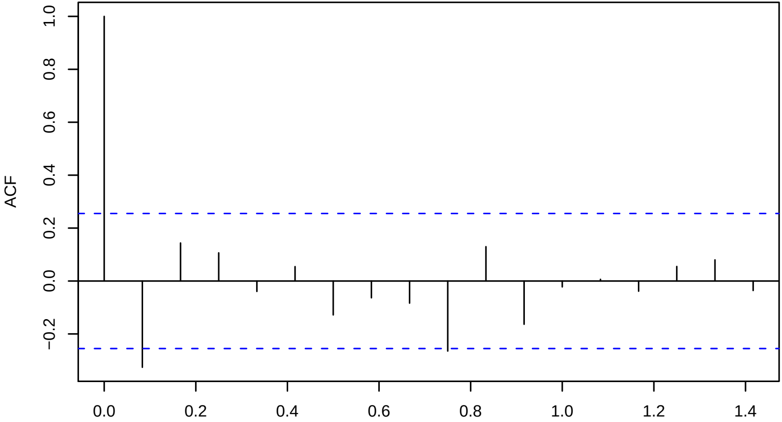
Error analysis for ICICI Bank

- Characteristics of ϵ.

	Y	ϵ
Mean	3.562	0.000
Std. Dev. 10.176		
Skewness	-0.718	0.106
	(5.162)	(0.113)
Kurtosis	3.975	3.442
	(2.377)	(0.488)
60 observations		

- Correlation between $(r_M, \epsilon) = -6.93e - 18$
- Serial dependence in ϵ: autocorrelation coefficient function of ϵ.

Serial dependence in errors



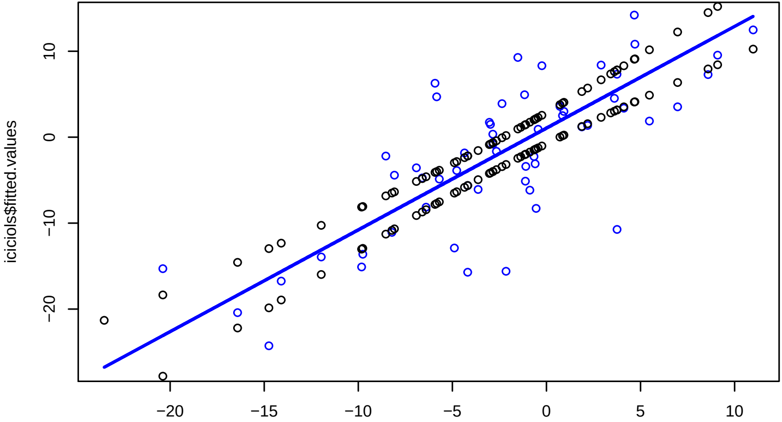
Model prediction for ICICI Bank returns

- $E(r_i - r_f) = \hat{\alpha} + \hat{\beta}_i E(r_M - r_f)$
- For a move in excess market index returns of 1%, what is the expected excess returns on ICICI Bank?
- $\hat{\beta}_i = 1.18$, $(r_M - r_f) = 1\%$
- $E(r_i - r_f) = 1.18 * 1 = 1.0513 + 1.18 = 2.2313\%$
- The 95% confidence interval: $0.3198 - 4.1502\%$
- Suppose we ignore the $\hat{\alpha}$.
- Then, E$(r_i - r_f) = 1.18\%$ with var$(E(r_i - r_f))$

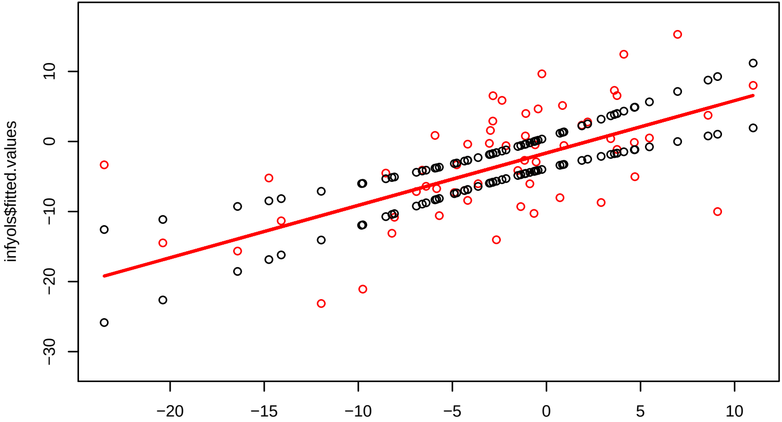
$$1 \times \mathrm{var}(\hat{\beta}_i) = \sigma_e / \sqrt{N} \sigma_X = 0.1215$$

- Then, the 95% confidence interval becomes: $0.9417 - 1.418$

Predictions and the 95% CI bands: ICICI Bank



Predictions and the 95% CI bands: Infosys Tech



HW: Model prediction for Infosys Technologies

- **HW**: For the same period data, what is the E($r_i - r_f$) for Infosys Technologies?
- What is the 95% CI for the prediction with and without the α?
- Find the monthly returns that Infosys Technologies actually obtained in January 2008.
  How well did the model's 95% prediction fare?

Transformation of the data and implication on the results

Changing the scale of the dependent variable

- We have the model: $Y_t = \alpha + \beta X_t + u_t$
- What if we accidentally create $Y_{t^*} = 1000 Y_t$ and run the regression instead?
- Impact: New equation on $Y_{t^*} = 1000 Y_t$

$$
\begin{aligned}
\Rightarrow Y_{t^*} &= 1000(\alpha + \beta X_t + u_t) \\
&= 1000\alpha + 1000\beta X_t + 1000 u_t \\
\Rightarrow Y_{t^*} &= \alpha\prime + \beta\prime_i X_t + u_{t^*} \\
\Rightarrow \hat{\alpha}\prime &= 1000\hat{\alpha} \\
\hat{\beta}\prime &= 1000\hat{\beta} \\
\hat{\sigma}_{u_t^*} &= 1000\sigma_{u_t}
\end{aligned}
$$

- The impact on the regression results:
  - Standard error of regression $\times 1000$
  - Standard error of regression co-efficient $\times 1000$
  - But t-stats, R^2, F-stats are not changed because they are ratios. Scale measures cancel out.
- **Conclusion:** When Y_t is changed by K, values of ALL required co-efficients by K.

Changing the scale of one independent variable

- Say we mistakenly transform $X'_t = \frac{X_t}{100}$
- Impact on the estimation:

$$\begin{aligned} Y_t &= \alpha + \beta X_t + u_t \\ &= \alpha + \beta_{100} \frac{X_t}{100} + u_t \\ &= \alpha + \beta' X'_t + u_t \end{aligned}$$

  - $\hat{\beta}' = 100\hat{\beta}$
  - Standard Error of $\hat{\beta}'$ also increases by 100
  - But t-stats remain unchanged.
- **Conclusion:**
  1. Only the transformed variable's coefficient is affected.
  2. t-stats, ESS, R^2, F all remain unchanged
- Important for data combinations with diverse ranges of scale: Some X might take real values from 0 to 1. Simultaneously others can take values above 10,000.

  Also useful to detect *errors* during data entry: the scale of estimated coefficients appear different from expected values.

Example: Market model for Infosys against $(r_m - r_f)/100$

- $Y_t = (r_{\text{infosys}} - r_f)$
- $X_t = (r_{\text{nifty}} - r_f)$
- $X_t^* = (r_{\text{nifty}} - r_f)/100$
- Statistical features:

	X_t	X_t^*
Mean	-2.6922	-0.0269
Std Dev	6.777	0.06777
Skewness	-0.703	-0.7033
Kurtosis	3.8276	3.8276

Regression results: $Y = f(X^*)$

	Infosys Tech			
	Estimate	Std. Error	Estimate	Std. Error
Intercept	-1.6413	1.0974	-1.50	0.1403
Market returns*	74.7377	15.1594	4.93	0.0000
F-stat(1, 57) = 90.05				
prob value = 2.5e-13				
R^2 = 0.2989				
Adj R^2 = 0.2866				
Residual Std. Err: 7.825				

Example: Market model for Infosys with $(r_{\text{infosys}} - r_f)/100$

- $Y_t = (r_{\text{infosys}} - r_f)$
- $Y_t^* = (r_{\text{infosys}} - r_f)/100$
- $X_t = (r_{\text{nifty}} - r_f)$
- Statistical features:

	Y_t	Y_t^*
Mean	-2.6922	-0.0269
Std Dev	6.777	0.06777
Skewness	-0.703	-0.7033
Kurtosis	3.8276	3.8276

Regression results: $Y = f(X^*)$

	Infosys Tech			
	Estimate	Std. Error	Estimate	Std. Error
Intercept	-1.6413	1.0974	-1.50	0.1403
Market returns*	74.7377	15.1594	4.93	0.0000
F-stat(1, 57) = 90.05				
prob value = 2.5e-13				
R^2 = 0.2989				
Adj R^2 = 0.2866				
Residual Std. Err: 7.825				