EViewsによる応用ファイナンス入門

第1章

統計分析

- 第1章で利用している次の用語について具体的に解説します。
  - 大数の法則
  - 中心極限定理
  - 最尤法
  - 単位根検定

モンテカルロシミュレーション

- 第1章ではt分布についてモンテカルロシミュレーションを用いて説明しました。
- 何故、乱数を繰り返し発生させることで統計量の確率分布について知ることができるのでしょうか?
- モンテカルロシミュレーションは中心極限定理の考え方を利用したものです。

二項分布

- 正しいコインを10回なげる
- 表がでたら1、裏なら0とする確率変数Xを考え、さらにその和をrとする

$$r = x_1 + x_2 + \cdots + x_{10}$$

- 表の出る確率p=0.5で10回の試行を行う時、x回表のでる確率は二項分布に従う

$$f_{10}(x) = \quad {}_{10}C_x (1/2)^{10},\ x = 0,1,2,\cdots,10$$

二項分布

- rの期待値と分散は

$$E(r) = np = 5,\quad V(r) = np(1-p)\text{=2.5}$$

- 割合r/nの期待値と分散は

$$E(r/n) = p = 0.5$$

$$V(r/n) = p(1-p)/n\text{=0.025}$$

しかし、実際にコインを使って実験を行うと理論値のズレが生じる

大数の法則

- 試行回数Nを増やしていくとズレは小さくなる

$$P(|r/n - 0.5| \leq \varepsilon) \to 1 \ \ (n \to \infty)$$

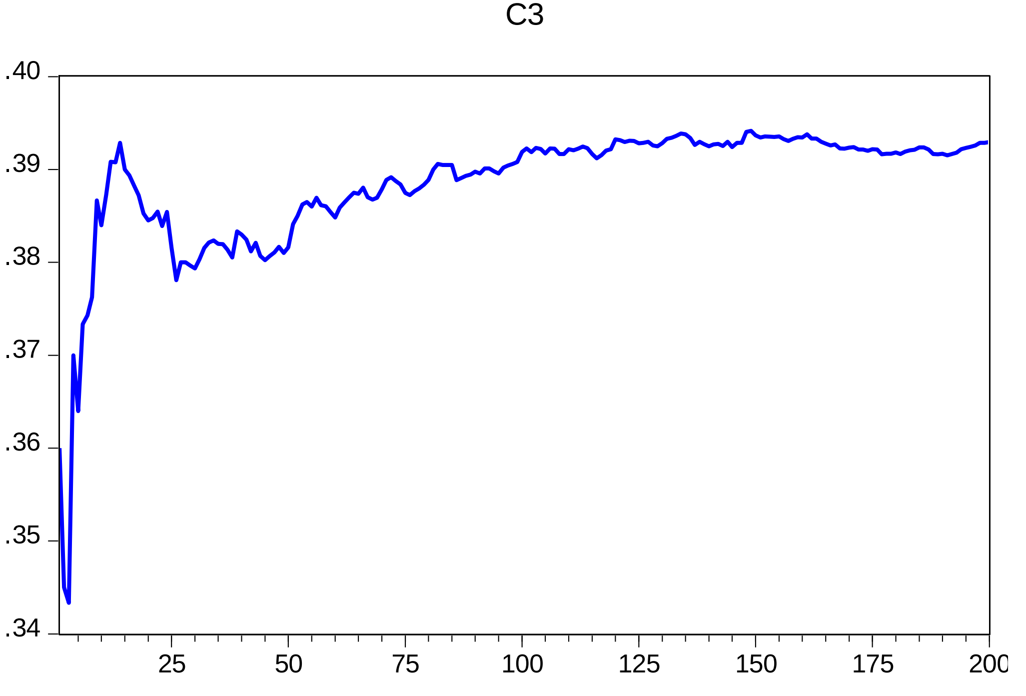
- 標本の大きさnが十分大ならば、標本平均の確率分布は母集団確率分布の平均(母平均)μの近くに集中していることを保証する

EViewsシミュレーション

- シミュレーション(統計学入門p157, 東京大学出版会)
  - $p = 0.6$であるベルヌーイ試行を2万回行う
  - 100回おきに区切り、1回目からの観測された成功率を計算する
  - 大数の法則によれば、この割合は試行回数nが増えるにつれて真の成功確率$p = 0.6$に近づく

EViewsシミュレーション

- 一様分布の乱数Uを発生させ、$U > 0.6$なら成功、$U \leq 0.6$なら失敗とする
- プログラムファイルはlln.prg



中心極限定理

- 母集団分布が何であっても、次に示す和の確率分布の形は、nが大きくなる時には大略正規分布と考えてよい

$$\bar{X} = (X_1 + \cdots + X_n)/n$$

- 母平均と母分散をそれぞれμとσ^2とすると、

$$S_n = X_1 + X_2 + \cdots + X_n \text{は} N(n\mu, n\sigma^2)$$

$$\bar{X} = (X_1 + X_2 + \cdots + X_n)/n \text{は} N(\mu, \sigma^2/n)$$

に従う

EViewsシミュレーション

- シミュレーション(統計学入門165, 東京大学出版会)
  - 二項分布$Bi(n,p)$に従う確率変数rを考える

$$r = x_1 + x_2 + \cdots + x_n$$

$x_1, x_2, \cdots, x_n$はそれぞれ独立にベルヌーイ分布
$Bi(1,p)$に従う

EViewsシミュレーション

ここで$\mu = E(x_i) = p, \sigma^2 = V(x_i) = p(1-p)$なので、
中心極限定理により、次の標準化係数は

$$z = \frac{(r - n\mu)}{\sqrt{n}\sigma} = (r - np)/\sqrt{np(1-p)}$$

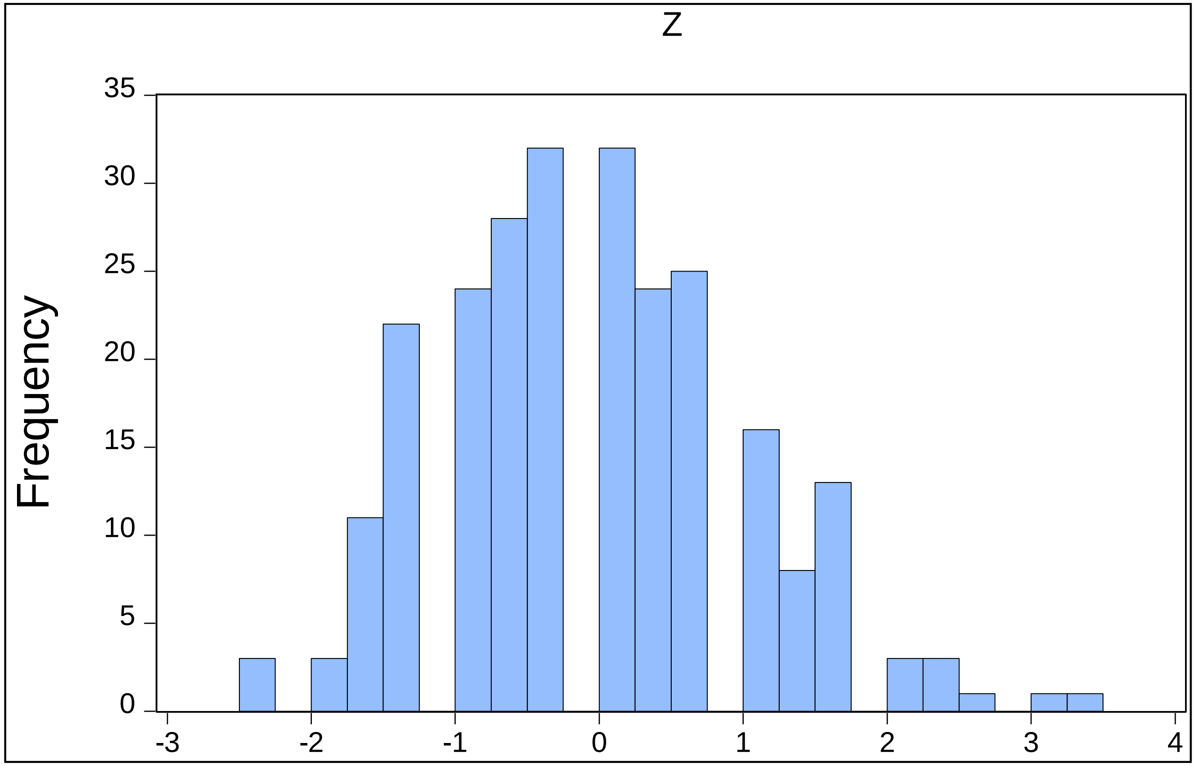
nが大きい時、標準正規分布N(0,1)に近づく。

EViewsシミュレーション

- シミュレーション
  - 成功確率$p = 0.1$($U < 0.1$の時成功、$U \geq 0.1$の時失敗)として$n = 5,10,20,30,50,100$に対して,二項乱数rを250個ずつ発生させる
  - それらを標準化し、分布を確認する
  - nが増えるにしたがって、標準正規分布に近づく
  - プログラム plt.prg

EViewsシミュレーション

- n=100の時の標準化変数の分布



最尤法

- 1を取る確率がp,0をとる確率が$1-p$のベルヌーイ分布$Bi(1,p)$を考える。ただし、$0<p<1$。
- 未知の母数pを推定したい
- 5回の試行で$X_1=1, X_2=0, X_3=0, X_4=1, X_5=1$の標本を得た
- 現実の標本は確率最大のものが実現したと仮定すると、この時の尤度は、

$$L(p)=p^3(1-p)^2$$

最尤法

- 例えば、$p = 0.1$と$p = 0.4$の場合を考えて尤度を計算する

$$p = 0.1 \quad \rightarrow \quad L(0.1) = 0.1^3(1-0.1)^2 = 0.00081$$

$$p = 0.6 \quad \rightarrow \quad L(0.6) = 0.6^3(1-0.6)^2 = 0.03456$$

$p = 0.4$の尤度が大きく、最もらしい

最尤法

- 実際に尤度関数を微分してみると、

$$\frac{dL(p)}{dp} = p^2(1-p)(3-5p)$$

$$p = 0.6$$

- 5回の試行において3回表がでているので、尤もらしい値といえる

尤度関数

- 未知の母数をθとすると、尤度関数は一般的に

$$L(\theta)=f(x_1,\theta)\cdot f(x_2,\theta)\cdot\cdots\cdot f(x_n,\theta)$$

となる。尤度はθの関数と考えることができる。

- 母集団分布が正規分布のように複数のパラメータから構成される場合は、

$$L(\theta_1,\theta_2,\cdots,\theta_k)=\prod_{i=1}^{n} f(x_i;\theta_1,\theta_2,\cdots,\theta_k)$$

となる。

尤度対数

- 尤度関数は積の形で、数学的な扱いが不便なので、対数をとって和の形にした対数尤度を利用する

$$\log L(\theta) = \sum_{i=1}^{n} log f(X_i, \theta)$$

最尤法推定

- 例えば、正規分布に従う確率変数Xの標本がn個ある時、母平均μと母分散σ^2を推定することを考える。
- 尤度関数は正規分布の密度関数を利用して、

$$L(\mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma} exp\{-(X_i - \mu)^2/2\sigma^2\}$$

$$\log L(\mu, \sigma^2) = -n \log\left(\sqrt{2\pi}\sigma\right) - \sum_{i=1}^{n}(X_i - \mu)^2/2\sigma^2$$

- EViewsのLogLオブジェクトに記述する時は、次のようにする

$$\log \frac{1}{\sqrt{2\pi}\sigma} exp\{-(X_i - \mu)^2/2\sigma^2\}$$

対数尤度の計算

- 第一章では最小二乗法で推定したモデルの対数尤度を次式で計算しています。

$$\mathrm{LL} = -\frac{T}{2}\left(1 + \log 2\pi + \log \frac{\hat{\epsilon}'\hat{\epsilon}}{T}\right)$$

- ここでは確率変数が残差であり、その期待値がゼロで残差の分散を次のように表現しています。

$$\sigma^2 = \frac{\hat{\epsilon}'\hat{\epsilon}}{T}$$

対数尤度関数の例

- この時、対数尤度関数LLは

$$
\begin{aligned}
LL &= \sum_{i=1}^{T} \log f(\epsilon) = \sum_{i=1}^{T} \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{\epsilon_i^2}{2\sigma^2}\right)\right) \\
&= \sum_{i=1}^{T} \log(2\pi\sigma^2)^{-\frac{1}{2}} + \sum_{i=1}^{T} \log\left(exp\left(-\frac{\epsilon_i^2}{2\sigma^2}\right)\right) \\
&= -\frac{T}{2}\log(2\pi\sigma^2) - \frac{\sum_{i=1}^{T}\epsilon_i^2}{2\sigma^2} \\
&= -\frac{T}{2}\log(2\pi) - \frac{T}{2}\log\left(\frac{\hat{\epsilon}'\hat{\epsilon}}{T}\right) - \frac{T}{2} \\
&= -\frac{T}{2}\left(1 + \log 2\pi + \log\frac{\hat{\epsilon}'\hat{\epsilon}}{T}\right)
\end{aligned}
$$

単位根検定

- 次に示すようなモデルを考えます。

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 trend + \epsilon_t$$

- 単位根検定はβ_1=1を仮説検定します。したがって、両辺からy_{t-1}を引いて、次のように変形します(Dickey-Fuller)。

$$\Delta y_t = \alpha + (\beta_1 - 1) y_{t-1} + \beta_2 trend + \epsilon_t$$

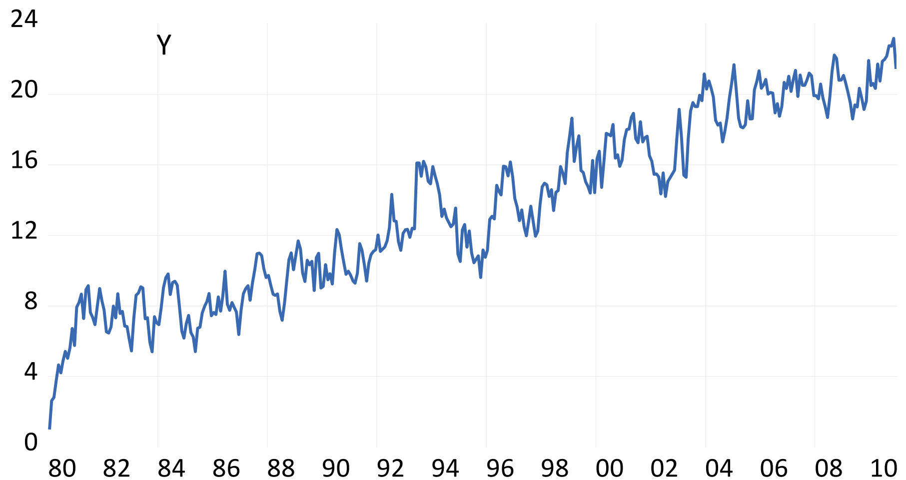
- さらにモデルの当てはまりをよくするために、Δy_tのラグ項を追加したものがAugmented Dickey-Fuller検定です。

単位根検定

- プログラムunit.prgで人工的なデータを作成します。

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 trend + \epsilon_t$$

- ここで$\alpha = 1, \beta_1 = 0.8, \beta_2 = 0.009, \epsilon_t$は標準正規分布に従う乱数とします。また、$y_{t-1}$=0とします。
- 作成したy_tのグラフ



単位根検定

- y_tでADFによる単位根検定(trend+intercept) を実行すると、$\beta_1 = 1$の帰無仮説は棄却されます。

Null Hypothesis: Y has a unit root
Exogenous: Constant, Linear Trend
Lag Length: 0 (Automatic - based on SIC, maxlag=16)

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-7.047006	0.0000
Test critical values:	1% level	-3.982988	
	5% level	-3.421983	
	10% level	-3.133816	

*MacKinnon (1996) one-sided p-values.

単位根検定

- 単位根検定の結果画面の下には次式の推定結果を表示します。

$$\Delta y_t = \alpha + (\beta_1 - 1)y_{t-1} + \beta_2 trend + \epsilon_t$$

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(Y)
Method: Least Squares
Date: 07/03/18 Time: 15:51
Sample (adjusted): 1980M02 2010M12
Included observations: 371 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
Y(-1)	-0.225518	0.032002	-7.047006	0.0000
C	1.264526	0.190343	6.643424	0.0000
@TREND("1980M01")	0.009885	0.001495	6.611427	0.0000

R-squared	0.119366	Mean dependent var	0.055089
Adjusted R-squared	0.114580	S.D. dependent var	0.946285
S.E. of regression	0.890423	Akaike info crite...	2.613814
Sum squared resid	291.7702	Schwarz criterion	2.645481
Log likelihood	-481.8625	Hannan-Quinn criter.	2.626391

$\beta_1 - 1$の値は約-0.23となっています。

単位根検定

- ダイアログの設定

$$y_t = \alpha + \beta_1 y_{t-1} + \beta_2 trend + \epsilon_t$$

グラフを表示して自分が想定したモデルに応じて、選択肢を選びます。