Itaú BBA

Dados os retornos $\{r_1, \dots, r_T\}$ e $H_0 : S(r) = 0$
$H_a : S(r) \neq 0$

estatística t de $\hat{S}(r)$: $t = \frac{\hat{S}(r)}{\sqrt{6/T}}$

regra de decisão: rejeitar H_0 num nível de signif. α se $|t| > Z_{\alpha/2}$

$Z_{\alpha/2}$: quantil $100\frac{\alpha}{2}$ superior de $N(0,1)$

Da mesma forma; $H_0 : K(r) - 3 = 0$ versus $H_a : K(r) - 3 \neq 0$

$t = \frac{\hat{K}(r) - 3}{\sqrt{24/T}}$ rejeitar H_0 se $|t| > Z_{\alpha/2}$

Teste de Jarque & Bera : combinação dos dois testes acima

estatística teste: $JB = \frac{\hat{S}^2(r)}{6/T} + \frac{(\hat{K}(r) - 3)^2}{24/T}$

cuja distribuição assintótica é $\chi^2(df = 2)$

H_0 : r_t é normal

Exemplo IBM

Itaú BBA

Distribuições de Retornos

O modelo mais geral para os log-retornos $\{r_{it};\ i=1,\dots,N;\ t=1,\dots,T\}$ é a sua função distribuição conjunta

(2) $$F_r(r_{11},\dots,r_{N1};\ r_{12},\dots,r_{N2};\ r_{1T},\dots,r_{NT};\ Y,\Theta)$$

na qual: Y é um "vetor de estado" composto das variáveis que descrevem o "ambiente" no qual os retornos são determinados

Θ é um vetor de parâmetros que determina univocamente a função de distribuição $F_r(\cdot)$

$F_r(\cdot)$ dita o comportamento estocástico de r_{it} e de Y

Em geral Y é tratado como dado e a principal ~~preocu~~ preocupação passa a ser a distribuição condicional de $\{r_{it}\}$ dado Y

Análise empírica: estimar Θ e dado o histórico passado dos log-retornos tirar conclusões, via inferência estatística, sobre o comportamento de $\{r_{it}\}$

O modelo da equação (2) é muito geral para ser de uso prático. Apesar disso ele fornece o arcabouço geral para modelos econométricos de $\{r_{it}\}$.

Exemplo: CAPM tem como foco a distribuição conjunta de N retornos em um único instante de tempo t, i.e., a distribuição de

$$\{r_{1t},\dots,r_{Nt}\}$$

Outras teorias enfatizam a estrutura dinâmica de retornos de ativos individuais, i.e., a distribuição de

$$\{r_{i1}, \dots, r_{iT}\} \quad \text{p/ um dado ativo } i$$

Dado o foco é a distrib. conjunta de $\{r_{it}\}_{t=1}^{T}$ p/ o ativo i (análise univariada) é útil particionar a distribuição conjunta como

(3) $$F(r_{i1}, \dots, r_{iT}; \Theta) = F(r_{i1}) F(r_{i2} | r_{i1}) \cdots F(r_{iT} | r_{i,T-1}, \dots, r_{i1})$$
$$= F(r_{i1}) \prod_{t=2}^{T} F(r_{i,t} | r_{i,t-1}, \dots, r_{i1})$$

(o vetor de parâmetros Θ foi omitido para simplificar)

Esta partição ressalta as <u>dependências temporais</u> dos log-retornos r_{it}.

O problema passa ser então a especificação da <u>distribuição condicional</u>

$$F(r_{it} | r_{i,t-1}, \dots, r_{i,1})$$

e, em particular, como esta evolui c/ o tempo.

<u>Na prática</u>: diferentes especificações p/ esta distribuição

$$\Downarrow$$

teorias diferentes

Itaú BBA

Exemplo: uma das versões da hipótese da caminhada aleatória assume que

$$F(r_{i,t} | r_{i,t-1}, \ldots, r_{i1}) = F(r_{i,t})$$

distrib. condicional ~ distribuição marginal

$\Longrightarrow$ retornos são temporalmente independentes e, portanto, não previsíveis.

Para retornos calculados em baixa-freqüência usa-se em geral um tratamento de variável aleatória contínua e, fazendo uso de sua p.d.f. e da relação (identid)

$$f_{x,y}(x,y;\theta) = f_{x|y}(x;\theta)\, f_y(y;\theta)$$

pode-se escrever a partição anterior (eq. (3)) como

$$f(r_{i1}, \ldots, r_{iT}; \Theta) = f(r_{i1}, \Theta) \prod_{t=2}^{T} f(r_{it} | r_{i,t-1}, \ldots, r_{i1}; \Theta)$$

Para retornos em alta-freqüência o caráter discreto passa a ser um problema. Tick-size importa

NYSE: até julho de 1997 tick-size = US$ 1/8
até janeiro de 2001 US$ 1/16
após US$ 0.01

Itaú BBA

Como a soma de um número finito de v.a.s normais i.i.d. é normal, então $r_t[K]$ também é normal sob esta hipótese.

Além disso: não existe limite inferior para r_t e o limite inferior para R_t é satisfeito utilizando-se

$$1+R_t = e^{r_t}$$

Entretanto: continua o problema do excesso de curtose positivo

Distribuições estáveis

- Generalizações naturais da distribuição normal (estáveis pelo somadas)
- capturam o excesso de curtose positiva dos retornos históricos empíricos.
- entretanto: distribuições estáveis não normais podem ter variância infinita. (conflito c/ teorias financeiras)
- além disso: modelagem estatística é mais difícil

Exemplo: distribuição de Cauchy (simétrica e c/ variância infinita)

Itaú BBA

Se $X_1, X_2, \ldots$ são va's independentes e identicamente distribuídas, com média μ e variância σ^2 então

$$(X_1 + \cdots + X_n - n\mu) / \sigma\sqrt{n}$$

converge em distribuição para uma v.a. c/ distrib. normal padrão.

Teorema limite: $X_1, \ldots, X_n$ i.i.d. então $\dfrac{\sum_{i=1}^{n} X_i}{A_n} - B_n \xrightarrow{d} X$

Quais as leis de limite que aparecem desta forma?

Suponhamos X uma v.a. e que, para cada n, existam constantes a_n, b_n tais que

$$a_n X + b_n \overset{d}{=} X_1 + X_2 + \cdots + X_n$$

na qual $X_1, \ldots, X_n$ são i.i.d. e c/ a mesma distrib. de X.

Neste caso dizemos que X é uma v.a. c/ distrib. estável

Exemplos: normal e Cauchy

Cauchy: $f(x) = \dfrac{1}{\pi} \dfrac{1}{\gamma^2 + (x - \delta)^2}$

localização

escala
$(\gamma > 0)$

Mistura de Distribuições Normais

Estudos recentes tem utilizado, para retornos de ações, as chamadas misturas de escala ou misturas finitas de distribuições gaussianas.

Sob a hipótese de que os retornos R_t são distribuídos de acordo c/ uma SMN os log-retornos $r_t \sim N(\mu, \sigma^2)$

Entretanto σ^2 é uma v.a. c/ distrib. positiva (e.g.

$\sigma^{-2} \sim \Gamma$

Exemplo : $r_t \sim (1-X)N(\mu, \sigma_1^2) + XN(\mu, \sigma_2^2)$

X : Bernoulli tal que $P(X=1) = \alpha$ $\quad 0 < \alpha < 1$

$P(X=0) = 1-\alpha$

$\sigma_1^2 < \sigma_2^2$

$\alpha = 0{,}05 \longrightarrow$ 95% dos retornos seguem $N(\mu, \sigma_1^2)$ e

5% $N(\mu, \sigma_2^2)$

$\sigma_2^2 > \sigma_1^2$ coloca mais "massa" nas caudas da distribuição dos retornos mas a maioria dos log retornos ainda seguem a $N(\mu, \sigma_1^2)$

Vantagens:
- mantém a tratabilidade da gaussiana.
- momentos de ordem superior finitos
- pode capturar o excesso de curtose.

Entretanto: . a estimação dos parametros ainda é complicada

Itaú BBA

Retornos Multivariados

Dado um vetor aleatório $X = (X_1, \ldots, X_p)$, ~~seu~~ seu vetor média e matriz de covariância são dados (definidos) como

$$E[X] = \mu_x = (E[X_1], \ldots, E[X_p])'$$

$$Cov[X] = \Sigma_x = E[(X-\mu_x)(X-\mu_x)']$$

uma vez que os valores esperados existam.

Quando uma amostra $\{x_1, \ldots, x_T\}$ está disponível a média e a covariância amostral são definidas como

$$\hat{\mu}_x = \frac{1}{T}\sum_{t=1}^{T} x_t \qquad \hat{\Sigma}_x = \frac{1}{T-1}\sum_{t=1}^{T} (x_t - \hat{\mu}_x)(x_t - \hat{\mu}_x)'$$

Estas estatísticas amostrais são consistentes uma vez que a matriz de covariância de X seja bem definida

Em finanças a distribuição normal multivariada é em geral utilizada para o log retorno r_t.

Função Verossimilhança dos retornos

A partição da equação (3) pode ser utilizada para obtermos a função de verossimilhança dos log retornos $\{r_1, \ldots, r_T\}$ p/ um ativo (i foi omitido!)

Se a distribuição condicional $f(r_t | r_{t-1}, \ldots, r_1; \Theta)$ é normal, c/ média μ_t e variância σ_t^2

Dist. Cjta. = Dist. Cond × Dist. Marg.

$f(x,y) = f(x|y) f(y)$

$\{r_1, r_2\}$ consecutivo.

$$f(r_1,r_2) = f(r_2|r_1) f(r_1)$$

$\{r_1, r_2, r_3\}$

$$f(r_1,r_2,r_3) = f(r_3|r_2,r_1) f(r_2,r_1)$$
$$= f(r_3|r_2,r_1) f(r_2|r_1) f(r_1)$$

$\vdots$

$\{r_T, r_{T-1}, \dots, r_1\}$

filtr

$$f(r_T, r_{T-1}, \dots, r_2, r_1) = \left\{\prod_{t=2}^{T} f(r_t | r_{t-1}, \dots, r_1)\right\} f(r_1)$$

Se $r_t | r_{t-1}, \dots, r_1 \sim N(\mu_t, \sigma_t^2)$

$$f(r_T, \dots, r_1) = \left\{\prod_{t=2}^{T} \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(r_t-\mu_t)^2}{2\sigma_t^2}}\right\} f(r_1)$$