

# 第4章 收益波动率计算

清华大学经管学院 朱世武

[Zhushw@sem.tsinghua.edu.cn](mailto:Zhushw@sem.tsinghua.edu.cn)

Resdat样本数据: [www.resset.cn](http://www.resset.cn)

SAS论坛: [www.resset.cn](http://www.resset.cn)

*The Power to Know.*

# 波动率估计法

收益波动率计算是金融计算与建模的基础，如风险度量、资产定价等。

最简单的收益波动率计算模型是静态波动率估计模型。实际中用的最成功、最常用的方法是移动平均、指数平滑和GARCH模型。

# 移动平均模型

表4.1 移动平均法估计波动率

等权重	指数加权
$\hat{\sigma} = \sqrt{\frac{1}{M-1} \sum_{t=1}^M (r_t - \bar{r})^2}$	$\hat{\sigma} = \sqrt{(1-\lambda) \sum_{t=1}^M \lambda^{t-1} (r_t - \bar{r})^2}$

注：近似公式  $\sum_{j=1}^T \lambda^{j-1} \cong \frac{1}{1-\lambda}$  。精确权重为  $\lambda^{j-1} / \sum_{j=1}^T \lambda^{j-1}$

**简单移动平均**（Simple Moving Average, SMA）模型是动态模型中最为简单的一种。它是以过去M天收益的样本方差来估计当前的波动率，即：

$$\sigma_t^2 = [1/(M-1)] \times \sum_{i=1}^M \left( r_{t-i} - \frac{\sum_{j=1}^M r_{t-j}}{M} \right)^2$$

这样每天通过增加前一天的信息和去掉第前M+1天的信息来更新预测。

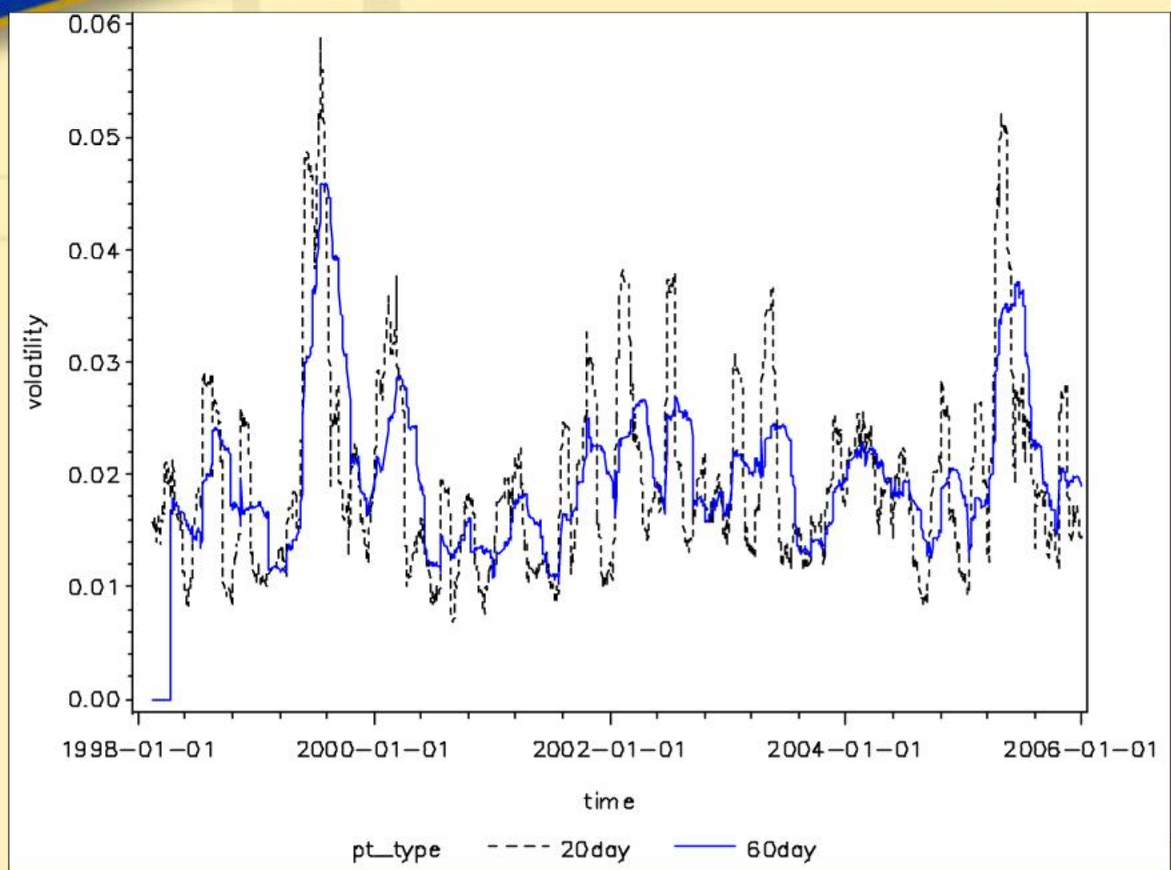


图4.1 波动率的时间曲线



指数加权移动平均模型依赖参数 $\lambda$ ，称 $\lambda$ 为衰减因子(decay factor)，该参数决定估计波动率时各观察数据的相对权重。

形式上，对 $t$ 时间波动率的预测为：

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda)[r_{t-1} - E(r_{t-1})]^2$$

其中，衰减因子 $\lambda$ 必须小于1。

当时间足够长时， $E(r_{t-1})$ 与 $E(r_t)$ 几乎相等。事实上，一般假设 $E(r_t)$ 约等于0，于是得到 $t$ 时刻波动率的如下预测：

$$\sigma_t^2 = \lambda \sigma_{t-1}^2 + (1 - \lambda)r_{t-1}^2$$

衰减因子 $\lambda$ 小于1。

对于日收益率数据，最优衰减因子  $\lambda$  为 0.94；对于月度收益率数据，最优衰减因子  $\lambda$  为 0.97。

# GARCH模型

GARCH (Generalized Autoregressive Conditional Heteroskedasticity) 模型称广义自回归条件异方差模型，或称为广义ARCH模型，

GARCH模型假定收益的方差服从一个可预测的过程，它依赖于最新的收益，也依赖于先前的方差。



GARCH (1, 1) 是这类模型中最简单的，用公式表示有：

$$r_t = \sqrt{h_t} e_t$$

$$h_t = \alpha_0 + \alpha_1 r_{t-1}^2 + \beta_1 h_{t-1}$$

$$e_t \sim iidN(0,1)$$

其中  $\alpha_0, \alpha_1$  和  $\beta_1$  均为待估的参数，可以用历史数据估计出。

# 波动率计算

## 计算环境

计算数据集：ResDat目录下的全部股票数据集，共30只。

需要宏文本文件：Stk.TXT。

时间区间：2005年。

计算日波动率；计算周、月或年波动率，可以用相应的收益率计算或直接由日波动率乘以一个相关因子。

对涨跌停板不作处理。

# 单个股票波动率计算

分别选择股票深发展（Stk000001）进行计算。时间区间为2005年。

日对数收益率计算：

```
options nodate nonotes nosource;  
data idxdate(keep=date);  
set ResDat.idx000001;  
where year(date)=2005; /*其他时间区间可修改此处*/  
%macro a(x);  
data a(keep=date r_1);  
set ResDat.&x;  
where year(date)=2005;  
r_1=log(mcfacpr*clpr)-log(lag(mcfacpr*clpr)); /*Mcfacpr为累积  
股价调整乘子 */
```

```
data log_ret(rename=(r_1=r&x));  
merge idxdate a;  
by date;  
if r_1=. then r_1=0;  
rr&x=r_1**2; /*日对数收益率的平方*/  
%mend a;  
%a(stk000001);  
run;
```

简单加权移动平均(**SMA**)计算的波动率:

```
proc sort data=log_ret;  
by Date;  
data sma(keep=Date);  
set log_ret;  
%macro a(x);  
data a;  
set log_ret;  
sum+rr&x;  
data b(keep=Date sma&x);  
merge a a (firstobs=21 rename=(sum=sum_1));  
sma&x=(sum_1-sum)/(20-1); /* 这里计算的是20天移动平均 */  
sma&x=sqrt(sma&x);  
proc sort data=b;  
by Date;
```



```
data sma;  
merge sma b;  
by Date;  
if sma&x=. then delete;  
%mend a;  
%a(Stk000001);  
run;
```

输出结果数据集**SMA**，包括变量有：

**DATE**：日期；

**SMAStk000001**：股票深发展收益日波动率。

指数加权(**EWMA**) 以及**GARCH(1,1)**计算的波动率留作练习  
(分别创建数据集**ewma**和**garch**)

*The Power to Know.*

## 三种模型结果比较

画出2005年，股票Stk000001（深发展）的对数收益图，图4.2。

```
%macro a(x);  
proc gplot data=log_ret;  
plot r&x*Date=1 ;  
symbol1 v=none i=join r=1 c=black line=1;  
%mend a;  
%a(Stk000001);  
run;
```

图4.2为深发展日收益时序图。

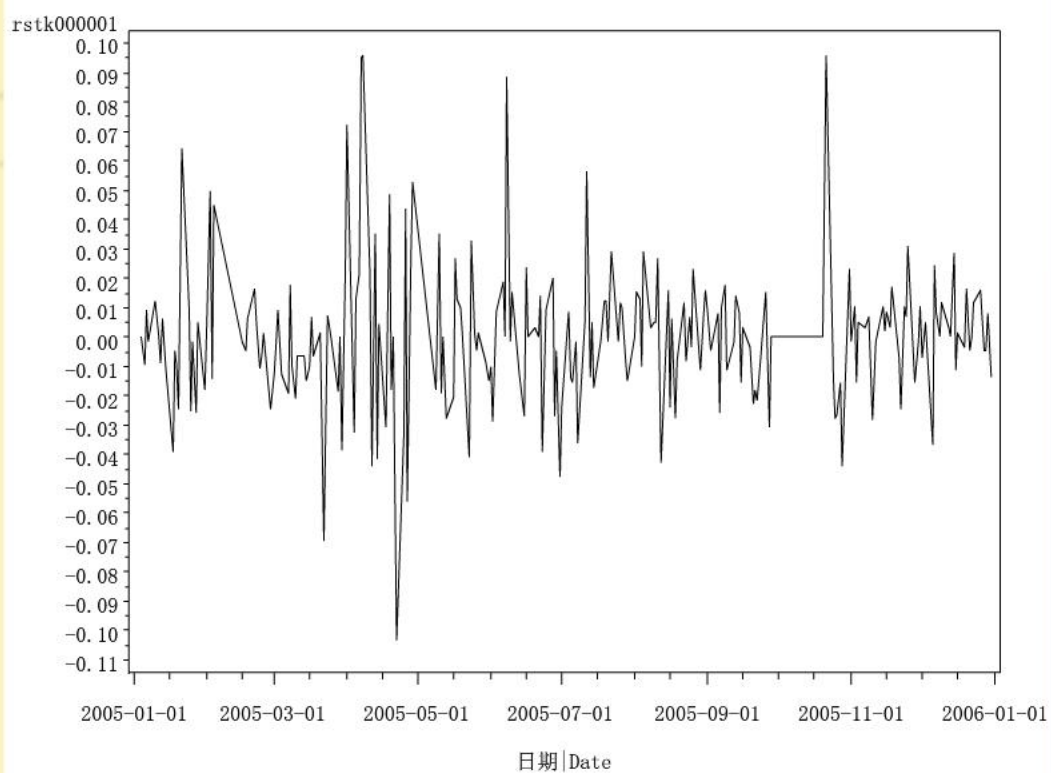


图4.2 深发展日收益时序图

三种模型求得的波动率时序图，图4.3——图4.6。

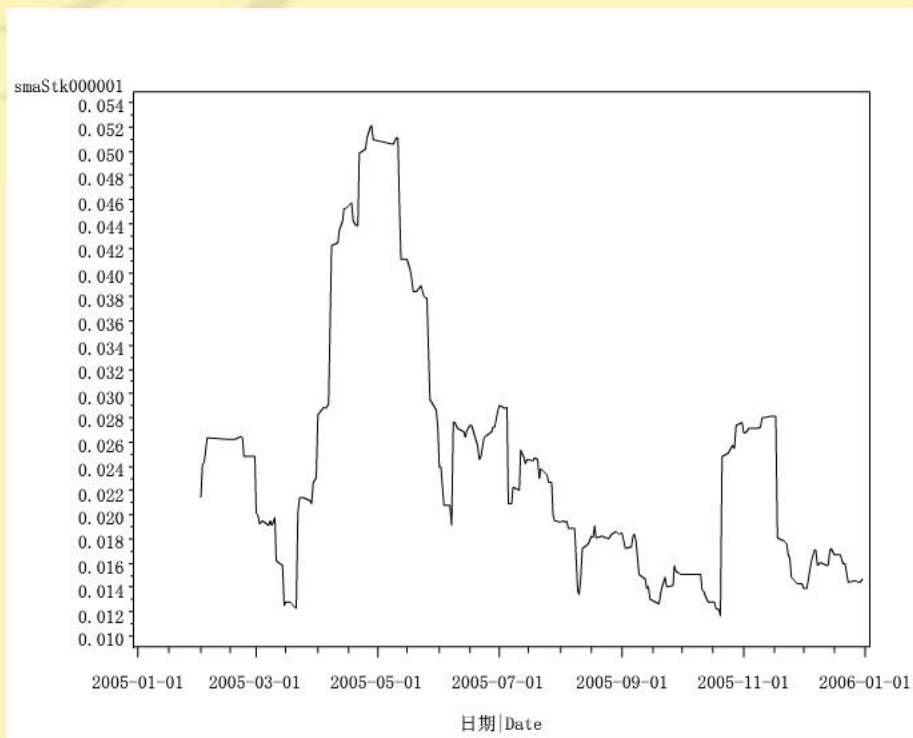


图4.3 简单加权移动平均(SMA)模型求得深发展日收益波动率时序图

*The Power to Know.*

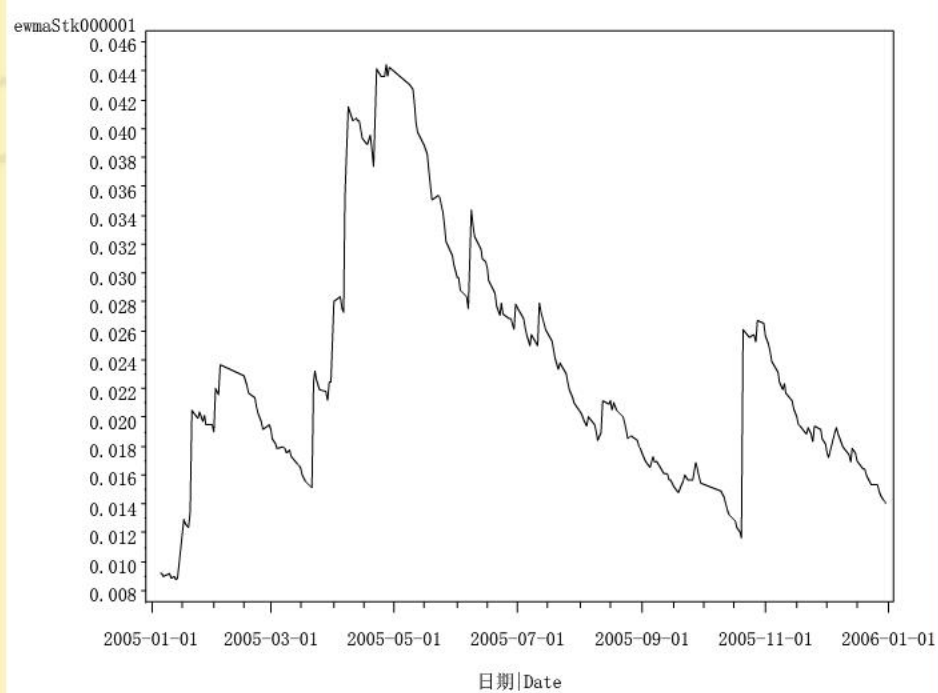


图4.4 指数加权(EWMA)模型求得深发展日收益波动率时序图



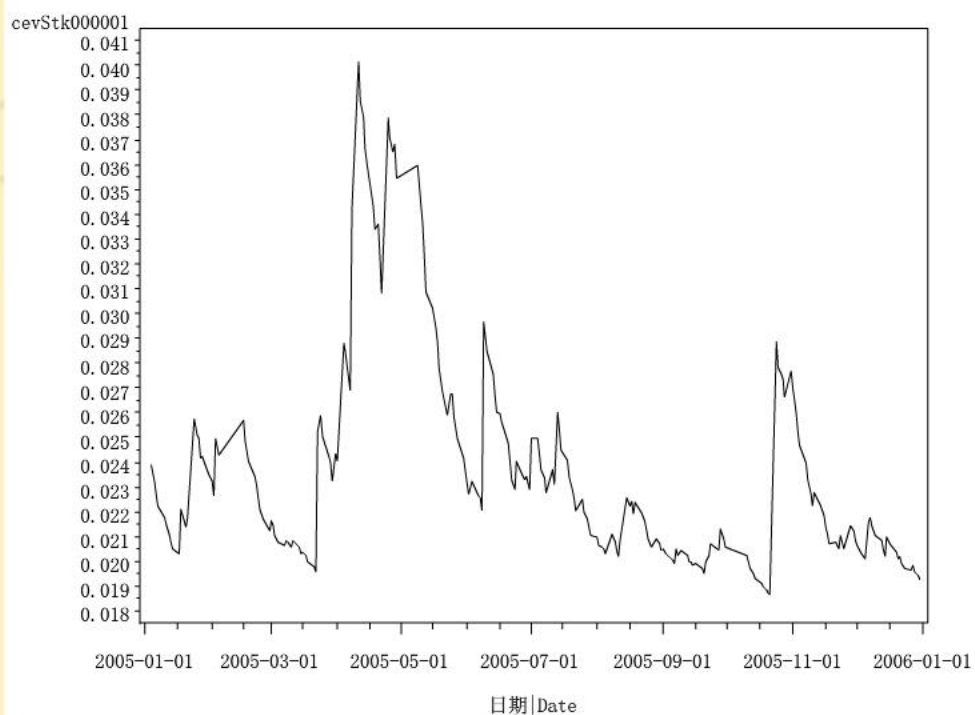


图4.5 GARCH(1,1)模型求得深发展日收益波动率时序图

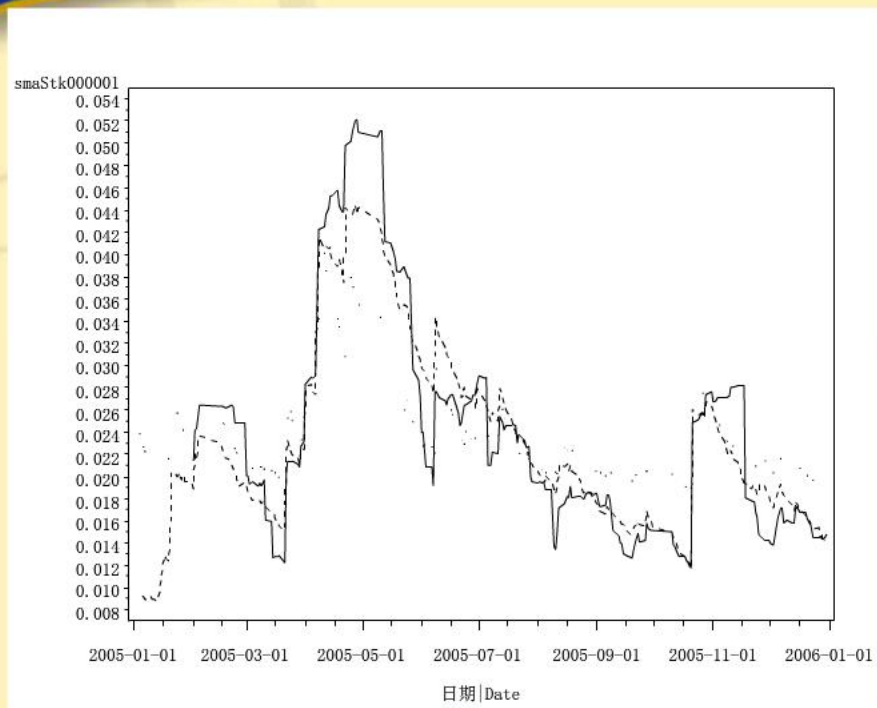


图4.6 三种模型求得深发展日收益波动率时序图

实线为简单加权移动平均（SMA）模型结果图，虚线为指数加权（EWMA）模型结果图，点线为GARCH(1,1)模型结果图。

## 多个股票波动率计算

利用简单加权移动平均（SMA）模型计算日波动率。

将日对数收益率5，10，20天交叠滚动，输出日波动率数据集：sma\_05；sma\_10；sma\_20。

上述三个数据集的变量包括：日期、用移动平均法计算的全部A股股票的日波动率。

计算上述三个数据集日波动率在全样本期的均值、最大、最小值，输出数据集：sma\_051；sma\_101；sma\_201。

上述三个数据集的变量包括：股票代码、股票名称、样本数、样本期波动率均值、最大值和最小值。

## 生成宏文本文件: **Stk.TXT**

```
data a;  
set ResDat.lstkinfo;  
a='%a(';  
b=',';  
c=')';  
file "Stk.txt" ;  
mac=a||'Stk'||stkcd||b||lstknm||c;  
put mac ;  
run;
```

## 计算程序（一个非常值得参考的程序）：

```
options nodate nonotes nosource;
filename mac "Stk.txt" ;
%macro b(aa,bb,cc);
data sma_&aa(keep=Date);
set ResDat.idx000001;
where 2001<=year(Date)<=2002; /*其他时间区间可修改此处*/
data sma_&bb;
set sma_&bb;
delete;
%macro a(x,y);
data a;
set ResDat.&x;
where 2001<=year(Date)<=2002; /*其他时间区间可修改此处*/
r_1=log(Mcfacpr*clpr)-log(lag(Mcfacpr)); /*Mcfacpr为累积股价调整乘子*/
data a;
set a nobs=nobs;
n=nobs;
id=1; /*为方便合并数据集所设的标志*/
```

*The Power to Know.*



```

/*保留起始日和结束日的股票价格和累积股价调整乘子，用来计算收益率均值*/
data b(keep=id adjclpr_begin adjclpr_end );
retain adjclpr_begin adjclpr_end ;
set a end=lastobs;
if _n_=1 then adjclpr_begin=clpr*Mcfacpr;
if lastobs then do;
adjclpr_end=clpr*Mcfacpr;
id=1;
output;
end;
data a(drop=adjclpr_begin adjclpr_end r_mean);
merge a b;
by id;
r_mean=(log(adjclpr_end)-log(adjclpr_begin))/n; /*收益率均值*/
r=r_1-r_mean;
rr=r*r;
data a;
set a;
sum+rr;
data b(keep=Date z&x);
merge a a (firstobs=&cc rename=(sum=sum_1));
z&x=(sum_1-sum)/(&aa-1); /*这里计算的是&aa天的移动平均*/
z&x=sqrt(z&x); /*用移动平均法计算的日波动率*/
if z&x=. then delete;

```

*The Power to Know.*

```
proc sort data=b;
by Date;
data sma_&aa; /*数据集中保存的是用移动平均法计算的日波动率*/
merge sma_&aa b;
by Date;
proc means data=b noprint;
var z&x;
output out=c max=max mean=mean min=min;
data c(drop=_type_);
set c;
stkcd=substr("&x",4,6);
lstknm="&y";
data sma_&bb; /*数据集中保存的是日波动率在全样本期的均值、最大、最小值*/
set sma_&bb c;
%mend a;
%include mac;
%mend b;
%b(05,051,6);
%b(10,101,11);
%b(20,201,21);
run;
```

# 等权重组合收益波动率

## 计算环境与输出数据集

样本数据集中**2005**年前上市的股票；**2005**年全年的日收益率。

计算数据集：ResDat.Lstkinfo, SAS逻辑库ResDat下所有股票数据集。

投资组合：随机抽股票组成投资组合，每个组合包括的股票数分别为：**2, 4, ..., 30**。每个组合各抽一次。

投资组合收益率：等权重收益率。

输出数据集：计算组合日收益率**2005**年的标准差。该过程可以重复**5**次，分别得到**5**个数据集，每个数据集包括的主要变量应该有：组合中股票的个数，该组合日收益率**2005**年的标准差。

作图：将组合中股票的个数与标准差作图，研究它们的关系，观测标准差有无随股票个数增加而减少的趋势。

*The Power to Know.*



# 实现算法

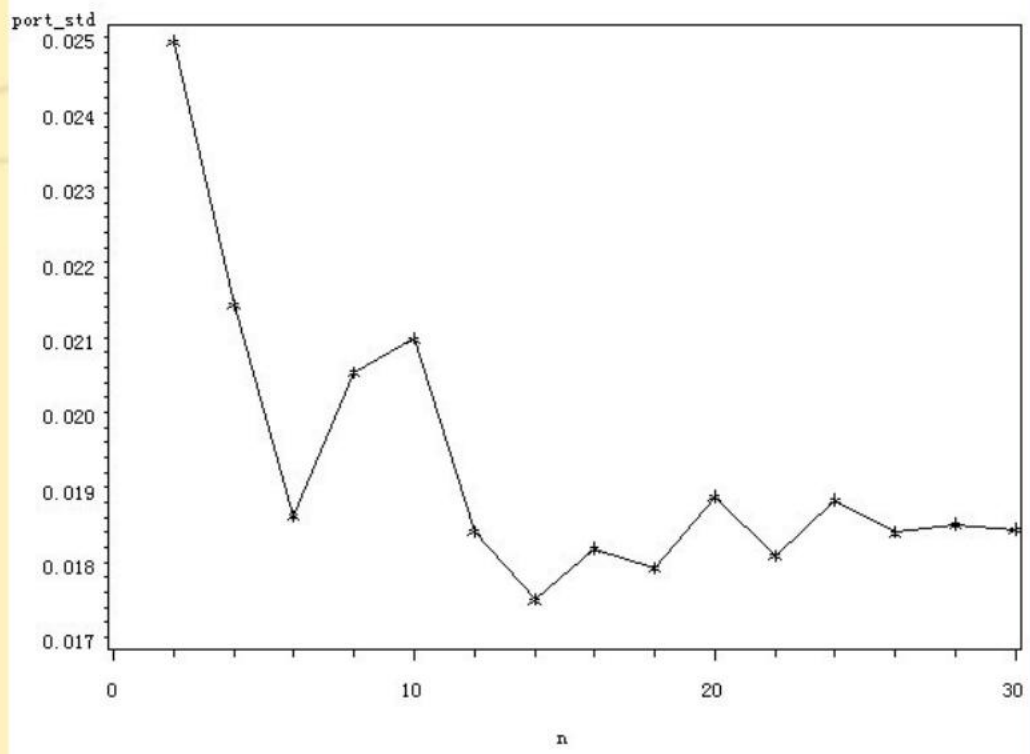
1. 挑选出2005年前上市的股票，这些股票的代码存于宏文本Stk2005.TXT，用于计算这些股票的收益率；
2. 对于第一步挑选的股票，计算2005年的日百分比收益，存于SAS数据集ResDat.R\_2005中；
3. 随机抽取一个投资组合，并求出相就的标准差；
4. 组合第三步的各段程序，用宏编写综合计算程序。

## 组合股票数与收益标准差二维图

改变综合计算程序中随机数产生器RANUNI(&I)中的种子，分别取值RANUNI (28668)，RANUNI(13878)得到以下3个组合股票数与收益标准差二维图形，见图4.7。

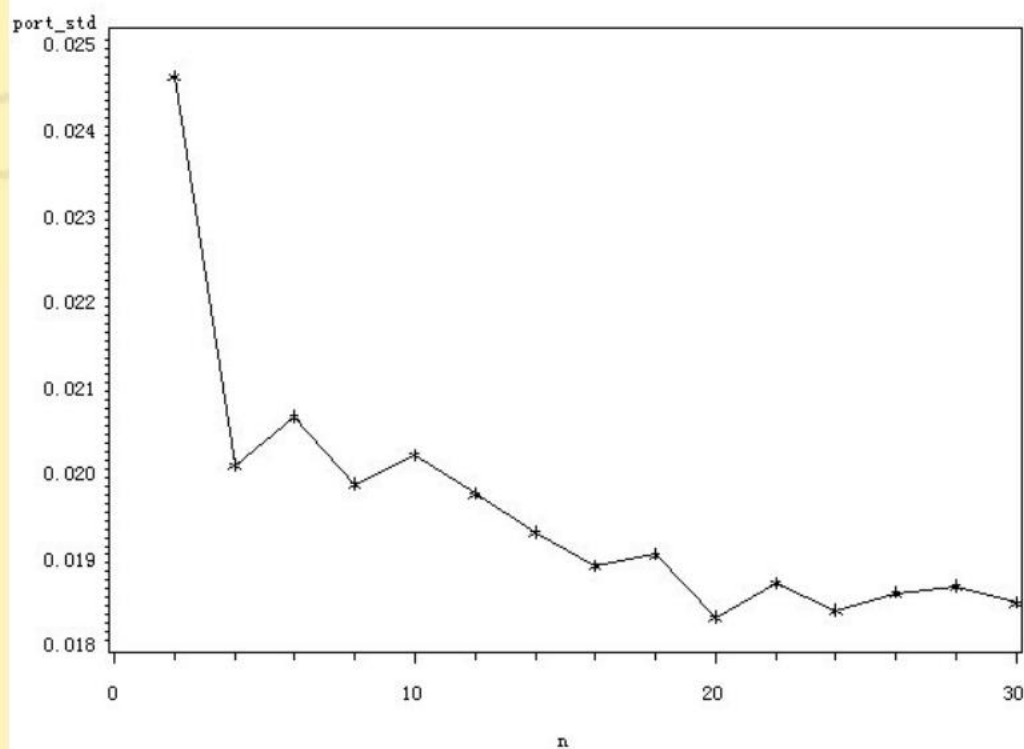
观测各图可以得到结论：当组合中的股票数到5、6支时，标准差得到显著的减少充分说明可以通过分散化投资组合来消除非系统风险。而当股票到20支以上时，继续增加组合中的股票数加标准差的减少并不明显。





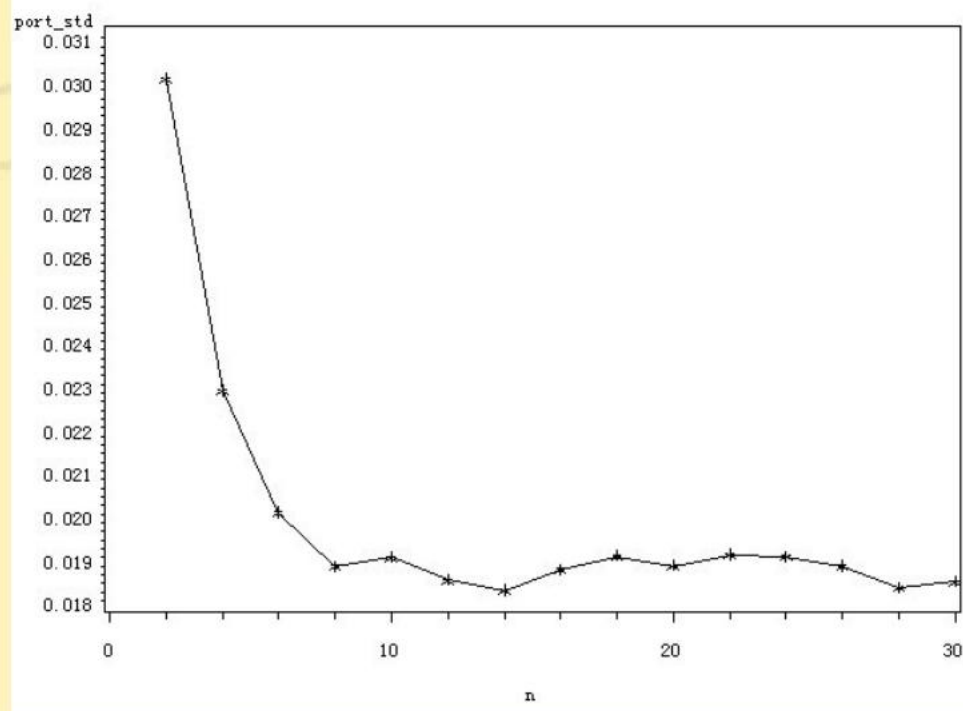
(1) **RANUNI (I)** (种子随组合中股票的个数而变化:  $I=2,4,\dots,30$ )

*The Power to Know.*



**(2) RANUNI (28668)**

*The Power to Know.*



**(3) RANUNI (13878)**