

国际金融试验班 2022 年秋 · 时间序列

# 第 1 讲：导论

授课人：刘 岩

武汉大学经管学院金融系

2022 年 9 月 5 日

# 本讲内容

① 课程介绍

② 导论

## 本节内容

① 课程介绍

② 导论

## 基本信息

- 课程主页：<http://www.liuyanecon.com/ug-ts-2022/>
- 电子邮箱：[ulysses1906@163.com](mailto:ulysses1906@163.com)
- 办公室：亮胜楼 B338-2
- 答疑时间：周四下午，邮件预约
- 授课周期：1-16 周，约 15 次课
- 本课程没有单一教科书，主要以上课讲授、课件为主

## 参考书：初级

\* 表示主要参考书

- ① Tsay, *An Introduction to Analysis of Financial Data with R*, 2013; 蔡瑞胸, 有教材配套网站, 包含大量 R 代码实例 (\*)
- ② Diebold, *Elements of Forecasting*, 4th edition, 2007; 中文版, 《经济预测基础教程》, 机械工业出版社, 2012 (\*)
- ③ Stock and Watson, *Introduction to Econometrics*, 3rd edition, 2011; 影印版, 《计量经济学》, 格致出版社, 2015 (\*)
  - 计量到时序的衔接, 第 2/3 章提供了很好的概率、统计基础知识概要
- ④ 何书元, 《应用时间序列分析》, 北京大学出版社, 2003

## 参考书：进阶

- ① Tsay, *Analysis of Financial Time Series*, 3rd edition, 2010; 中文翻译版, 《金融时间序列分析》, 人民邮电出版社, 2012
- ② Hayashi, *Econometrics*, Princeton University Press, 2000
- ③ Hamilton, *Time Series Analysis*, Princeton University Press, 1994; 中文翻译版, 《时间序列分析》, 中国人民大学出版社, 2015

## 成绩评定

- 平时成绩：10次作业，每次5%，共计50%
- 期末考试：考试周，闭卷（无x页纸小抄），占期末总成绩的50%
- 作业中的抄袭和考试中的作弊会受到严肃处理
- 留学生的要求与其他同学一致
- 不接受本学期出国交换的同学选修

## 助教安排及作业要求

- 刘芳瑞, 1349500819@qq.com
- 王汇轩, 1422869566@qq.com
- 作业会包括理论推导和（或）编程计算，编程作业需提交电子版答案
- 有编程内容的作业，助教会逐一运行大家所写代码
- 建议大家使用 L<sup>A</sup>T<sub>E</sub>X 编写作业答案：专业！
  - 推荐使用 T<sub>E</sub>XLive 2022 套装；X<sub>Y</sub>L<sup>A</sup>T<sub>E</sub>X 有完美中文支持
  - 课件生成：BEAMER + C<sub>T</sub>E<sub>X</sub> 中文宏包 + newpx 字体包



## 课程沟通

- 每次上课前一天，最新的课件会上传到课程主页
- 重要事项通过邮件进行沟通
- 信息登记：



## 邮件规范

- ① 请关闭各自邮箱的自动回复功能
- ② 有简短明确的主题
- ③ 有开头称谓，如 XX 老师，XX 助教
- ④ 正文另起一行，包括个人专业、班级等基本信息，并清楚陈述邮件内容
- ⑤ 落款，在初次邮件沟通时务必留全名
- ⑥ 电子邮件无需段首空格

## 邮件示例

【主题】22 时间序列课程：关于课件的问题

XX 老师：

你好！我是 XX 级 XX 专业学生，关于 XX 次课件 XX 页有 XX 问题，……

祝好！XX

## 软件

- 课程主要使用 R 进行教学示例以及完成编程作业
  - 可从下列地址下载 R 安装包: <https://mirrors.tuna.tsinghua.edu.cn/CRAN/>
  - 同时下载集成编译套装 R-Studio: <https://www.rstudio.com/>
  - R 支持将程序和图、表结果同步输出到 L<sup>A</sup>T<sub>E</sub>X 文档
- 可能有部分内容会涉及到 MATLAB
- 推荐大家学习、使用 Python
- 本课程没有专门的实验课，通过作业进行编程实践
  - 设备问题请及时联系我或者助教

## R-Studio 示例

The screenshot displays the RStudio environment with the following components:

- Source Editor:** Contains R code for reading an Excel file and creating a data frame. The code includes comments in Chinese and uses functions like `read_excel`, `data.frame`, `paste`, `levels`, `matrix`, and `data.frame`.
- Environment:** Shows "Global Environment" and "Environment is empty".
- Files:** Shows the current file "R: Convert Character Vector between Encodings".
- Viewer:** Displays the R documentation for the `iconv` function, including a description, usage, arguments, and a table of argument details.
- Terminal:** Shows the R startup message, including the platform information (x86\_64-w64-mingw32/x64) and the license information.

```

217     fileEncoding = "GBK")
218
219 ## bank size data
220 Z <- read_excel("银行规模20180504.xlsx",
221               sheet = 1,
222               col_names = T)
223 Z <- data.frame(Z)
224 Z$year <- paste("y", Z$year, sep = "")
225 Zcode <- Z[,3]
226 Zcode <- levels(factor(Zcode)) # with whole country total
227 Zbank <- Zcode[Zcode != "w0000"]
228 Y <- matrix(nrow = length(Zcode),
229            ncol = nyear,
230            dimnames = list(c(1:length(Zcode)), year))
231 ZA <- data.frame(code = Zcode, Y) # asset
232 ZB <- ZA # liability
233 ZL <- ZA # loan
234 ZD <- ZA # deposit
235 for (i in 1:length(Zcode)){
236   <
237 }
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

**Terminal Output:**

```

Platform: x86_64-w64-mingw32/x64 (64-bit)

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

> |

```

**Documentation: Convert Character Vector between Encodings**

**Description**  
This uses system facilities to convert a character vector between encodings: the 'i' stands for 'internationalization'.

**Usage**  
`iconv(x, from = "", to = "", sub = NA, mark = TRUE, toRaw = FALSE)`  
`iconvlist()`

**Arguments**

Argument	Description
<code>x</code>	A character vector, or an object to be converted to a character vector by <code>as.character</code> , or a list with <code>NULL</code> and raw elements as returned by <code>iconv(toRaw = TRUE)</code> .
<code>from</code>	A character string describing the current encoding.
<code>to</code>	A character string describing the target encoding.
<code>sub</code>	character string. If not <code>NA</code> it is used to replace any non-convertible bytes in the input. (This would normally be a single character, but can be more.) If "byte" the indication is "<xx>" with the hex code of the byte.

## 课程内容

- 1 导论，概率论复习
- 2 统计复习
- 3 时间序列数据特征
- 4 平稳时间序列的周期性与谱
- 5 国庆假期
- 6 ARMA 过程的表示与性质
- 7 自回归模型的估计
- 8 自回归模型的推断
- 9 时间序列回归分析拓展
- 10 时间序列的预测
- 11 向量自回归模型理论
- 12 向量自回归模型应用
- 13 随机波动率与条件异方差模型
- 14 非平稳时间序列与单位根过程
- 15 协整与向量误差修正模型
- 16 状态空间模型

## 数学基础

### 随堂说明、强化

- 微积分：极限，求导，极值，积分
- 线性代数：矩阵运算，行列式，特征值与特征向量，对称矩阵的特征值分解
  - 推荐参考书：李尚志，《线性代数》（非数学专业），高等教育出版社，2011

### 2次课进行概率统计复习、梳理

- 概率论：随机变量，独立性，概率分布、密度，矩，条件概率与期望，正态分布，大数定律与中心极限定理
  - 很多内容与随机过程理论紧密相连
- 统计：总体与样本，假设检验，统计量，极大似然估计，矩估计，一致性，渐进分布

## 课程目标

### 理论与实践紧密结合

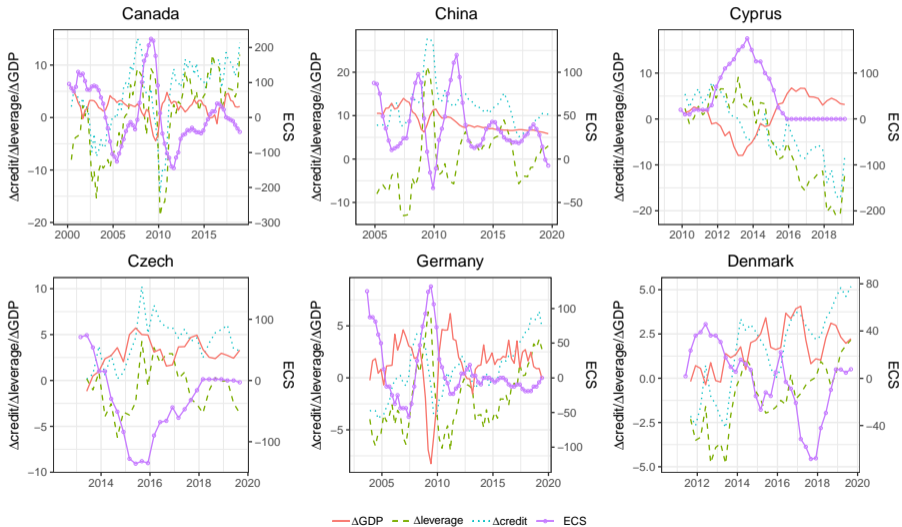
- 时间序列理论的基本概念、理论
- 常见时间序列模型的基本性质和估计、推断方法
- 掌握时间序列建模的基本步骤
- 能够用 R 软件实现对常见时间序列模型的实证分析
  - 实践使用场景：宏观研究，行业研究，金融量化分析



## 学习要点

- 不要怕！
  - 数学不是背公式，是“背”——理解——逻辑
  - 多接触，多“背”，就明白了
  - 碰到问题，多查 Wikipedia，部分知乎也还可以
- 重要的是原理，本科阶段不用纠结技术细节
- 会用：不但能把数据扔到 R 里跑出结果，还要看得懂结果、知晓原理
- 注意各个时序模型与实际经济、金融问题的联系
- R 编程就是学习怎么操作数据表，学会使用现成的宏包，正确设定模型，会画图（很多学问）

## R 专业绘图示例



## 本节内容

① 课程介绍

② 导论

## 时间序列：分析对象

- 时间序列：time series
- 广义而言，时间序列研究的就是随时间变化的一个随机序列
  - 随时间变化的随机序列就是随机过程：stochastic process
- 更具体的说法，时间序列理论关注的是随机过程  $\{X_t\}$  在时间上的相关性特征： $\text{cov}(X_t, X_{t-i})$ 
  - 随机过程理论的其他关注点包括：随机动态的概率特征（首达时），遍历特征（常返性），随机单调性（鞅），极限分布（ $t \rightarrow \infty$  时  $X_t$  的概率分布），随机收敛性（收敛到极限分布的速度）
- 时间上的相关性：对序列的动态行为进行统计推断与预测

## 时间序列：概率论视角

- 时间序列最早是概率论的一个分支：主要关心随机过程的平稳性与遍历性
  - 平稳性：对任意  $k$ ,  $(X_{t+k}, \dots, X_t)$  的联合分布或矩 (moments) 不随  $t$  变化
  - 遍历性：从任意初始值  $x_0$  出发,  $X_t$  随时间变化过程中可以取遍所有可能取值
- 这种一般（抽象）意义上概率性质，主要用途是统计物理中的理论性质：时间无限，空间无限
- 现实观测到的时间序列数据，时间、空间（取值）上均有限
- 有限样本的信息提取：概率论  $\implies$  统计学
  - 但统计学需要具体的应用

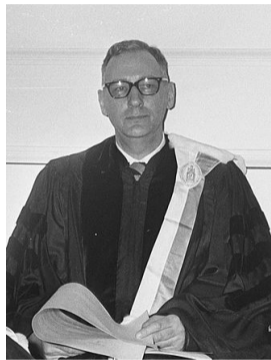
## 时间序列：经济的视角 I

- 熊彼特：数学天生适用于经济学分析，因为经济现象天然可以用数字描述
  - 经济变量随时间变化：形成大量的时间序列数据
  - 经济时间序列统计分析的开始：Cowles Foundation
    - 最初在 Chicago University, 50 年代搬到 Yale University
    - 主要是欧洲移民的数理经济学家，以 Koopmans 为代表
  - 受 Keynes 经济学影响，致力于用大型方程组来解释经济动态……
    - 多部门、动态 IS+LM+Phillips 曲线，数千个变量
- ……但并未充分意识到经济时间序列本身的动态性质

## 时间序列：经济的视角 II

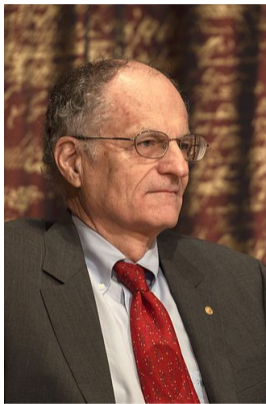
- 70 年代开始，“理性预期”革命彻底否定了传统 Keynes 经济学的研究方案，同时提出了以精简的时间序列模型为基础的新宏观经济学
- 代表人物：Tom Sargent 和 Chris Sims，2011 年诺奖得主
- 低维时间序列模型就可以很好的刻画经济的动态运行：产出、货币供给、利率、通胀、消费、投资、就业
- 原因就在于经济时间序列数据有典型的时间相关性特征
  - 单变量时间序列模型，向量自回归模型 (VAR)

## Koopmans, Sargent, & Sims



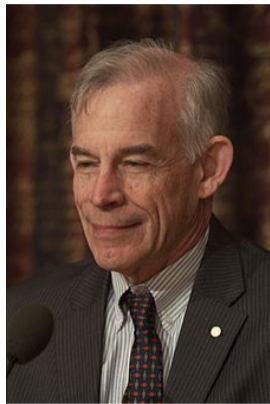
1910–1985

1975



1943–

2011



1942–

2011



## 时间序列：金融的视角 I

- 60年代，Chicago University, Center for Research on Security Prices (CRSP) 的建立，奠定了现代资产定价理论实证研究的基础
- 60-70年代最重要的金融理论：有效市场假说（弱、半强、强）
- 弱有效市场假说：历史收益率无法预测未来的收益率——当前的市场价格已经反映了所有公开市场信息

$$\mathbb{E}_t[r_{t+1} - r_t | r_t, r_{t-1}, \dots] = 0,$$

这意味着  $r_{t+1} - r_t$  与  $r_{t-i}$  之间的相关性为 0

- 时间序列提供了检验有效市场假说的理论与方法

## 时间序列：金融的视角 II

- Eugene Fama 和 Bob Shiller 2013 年因为实证资产定价的研究获得诺贝尔奖
- Fama 一直坚持“有效”市场的阵地
- Shiller 在 70 年代末到 80 年代广泛分析各种资产的收益-波动关系，由此认为市场是“无效”的，原因在于人的心理特征导致行为偏差，进而提出行为金融学

## Fama & Shiller



1939–

2013



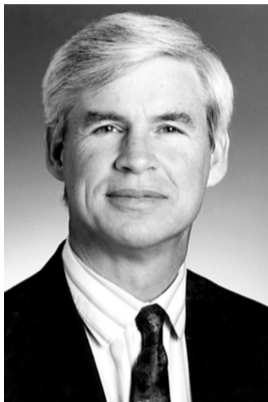
1946–

2013

## 时间序列：技术的进步

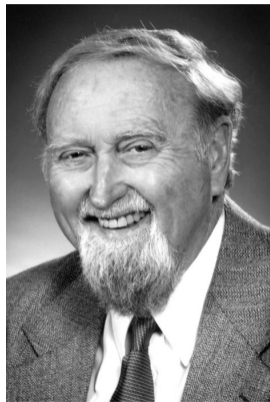
- 从 70 年代开始，时间序列逐渐成为经济与金融实证研究使用最广泛的基础性工具之一
- 而经济与金融研究的需要，也刺激了时间序列统计理论的发展——非平稳时间序列模型
- 代表人物：Bob Engel 和 Clive Granger，2003 年诺贝尔奖
  - 在单位根过程上发展起来的协整模型，用以描述多个非平稳序列间的长期均衡关系
  - 随机波动率模型，描述金融市场波动与回报之间的关系

## Engle & Granger



1942–

2003



1934–2004

2003

## 最近的新发展：机器学习

### 三大类机器学习 (machine learning) 算法

- 无监督学习 (unsupervised learning)
  - 目的：将一组数据进行聚类、拆分，使得每个数据子集具有一定的特征，最好是对应到人类语言的特定概念
  - 典型应用：图像识别，语音识别
- 监督学习 (supervised learning)
  - 有明确的推测目标  $Y$ ，要拟合  $Y$  与推测变量  $X$  间的函数  $Y = f(X)$ ，本质是函数拟合问题
  - 经济、金融中有大量应用  $\Rightarrow$  金融科技的主要方法支撑
- 强化学习 (reinforcement learning)
  - 解决高维的动态序贯决策问题，和动态规划 (dynamic programming) 紧密联系
  - 如 AlphaGo

# 跨国增长实证研究的模型不确定性问题： 机器学习的视角

刘岩，谢天



# 机器学习驱动的基本面量化投资研究

李 斌， 邵新月， 李玥阳