



2024-2025 学年度数据分析期末大作业

Term project of data analysis in 2024-2025 school year

报告题目：

日本劳动者在工作生活平衡中的性别差异

姓名：周毅荣

学号：22011125

指导老师：沈思

摘要

本研究旨在探讨日本劳动者在工作生活平衡中的性别差异。通过应用非参数检验方法，我们对不同性别、家庭角色、工时和改变工时意愿的关系进行了深入分析。

首先，我们对数据进行了基本描述性统计和初步分析，揭示了不同性别和家庭角色在工作时长及改变工作时间意愿上的显著差异。

接着，利用单一样本和两样本数据的位置和尺度推断方法，进一步分析了各类数据的分布特征和显著性差异。通过Kruskal-Wallis单因素方差分析和Spearman检验，我们发现了不同类别劳动者之间的显著差异。

然后，我们对月平均工作时间进行了线性回归分析，包含正态性检验、t检验、方差齐性检验、Durbin-Watson检验等基本步骤，便于与ARIMA模型进行对比。

最后，我们还分别建立了ARIMA模型和SARIMA模型对未来工作时间的变动进行预测。时间序列分析显示日本劳动者的月均工作时长存在显著的季节性波动，特别是在春季，这与日本的春季劳工运动有关。

综上数据分析结果得出结论，认为男性的工作时间减少速率远大于女性，在未来5年中将得到更好的工作生活平衡，而女性在工作中的地位也在提升。

关键词：工作生活平衡 性别差异 时间序列分析

ARIMA模型 SARIMA模型

目录

- 一、引言..... 5
 - 1.1 工作生活平衡的当代困境.....5
 - 1.2 日本社会的性别平等发展历史.....5
 - 1.3 日本工作环境中的加班文化.....6
 - 1.4 提出研究问题8
- 二、文献综述..... 8
 - 2.1 工作生活平衡的影响因素8
 - 2.2 相关研究的综述9
 - 2.3 当前研究的创新点 10
 - 2.4 数据的来源和介绍..... 10
- 三、性别和家庭角色的非参数检验.....11
 - 3.1. 描述性统计分析 11
 - 3.2 相关性分析 11
 - 3.2.1 分析概述 11
 - 3.2.2 卡方检验 12
 - 3.3 可视化分析..... 14
- 四、月平均工作时长的时间序列分析.....17
 - 4.1 数据的预处理..... 17
 - 4.1.1 描述性统计 17
 - 4.1.2 正态性检验 17
 - 4.1.3 方差齐性检验 18
 - 4.1.4 t检验 19
 - 4.2 时间序列的回归分析..... 19
 - 4.2.1 假设检验 19
 - 4.2.2 系数检验 19
 - 4.2.3 残差分析 20
 - 4.2.4 结果解读..... 21
 - 4.3 基于ARIMA模型的时间序列分析 21
 - 4.3.1 ARIMA模型概况..... 21
 - 4.3.2 实验环境..... 22

4.3.3	平稳性检验	23
4.3.4	白噪声检验	24
4.3.5	季节性分析	25
4.3.6	模型估计	26
4.3.7	模型拟合	26
4.3.7.1	ARIMA模型	26
4.3.7.2	SARIMA模型	27
4.3.8	残差白噪声检验	28
4.3.9	模型评估和预测	29
五、结果的实际意义与建议		30
5.1	实际意义	30
5.2	合理化建议	30
六、参考文献		31
七、附录		31
7.1	ARIMA模型拟合度和残差数据	33
7.2	SARIMA模型拟合度和残差数据	34
7.3	寻找模型参数的差分检验代码示例	34
八、致谢		36

一、引言

1.1 工作生活平衡的当代困境

工作生活平衡（Work-life Balance, 简称WLB），这一术语最早出现在 80 年代的英国，是当时妇女解放运动的一项纲领。该运动提倡女性享有灵活的工作时间安排和产假。但是，尽管男性在社会上不受束缚，可以追求自己的职业目标而不必担心家务和抚养孩子，但职场女性却被期望在工作的同时承担家务和抚养孩子的责任。在 80 年代，经常有人指出这种明显的工作与生活不平衡，并质问职场女性是否真的可以“拥有一切”。

工作时长，工作安排方式，以及是否有休息时间，不仅会大大影响工作质量，也会影响到工作时间以外的生活。对工作时间问题的重视已经成为了《国际劳工组织章程》的依据。作为《国际劳工组织章程》基础的《凡尔赛和约》第十三章明确承认，对工作时间的管理，包括确定每天和每周的最高工作时长，是改善当时普遍存在的劳动条件所迫切需要的措施之一。^[1]

然而，数字时代模糊了工作与生活的界限，威胁到弹性工作制所带来的自由。当“下班不回老板信息”变成一种奢侈，员工们经常会被迫利用弹性工作制来保持与工作的数字连接。我们越来越难以知道工作何时结束、生活何时开始。而且，人类越精通互联网技术，就越难在两者之间划清界限。

所以，员工该如何为职业和个人活动设定界限变得很重要，这将确保两种生活不会发生冲突。^[2]由于技术和我们的“永远在线”文化不会很快消失，因此关于工作与生活平衡的讨论需要不断发展，并且已经成为了当今国际社会的重要议题。

1.2 日本社会的性别平等发展历史

在江户时代的近代日本，由于儒家统治阶级的强化，男女的统治地位更加强大，但对此，神道教的三山益保说：“人的生命之根本在于”男女和谐。”这与日本这个和谐国家的传统神道仪式和民间生育仪式有关，在性别和谐的世界里，男女是平等的。”他批评了这种情况其中，以家庭为基础的婚姻制度疏远了男女之间的“爱的感觉”（佐佐木纯之介等人，《日本历史概论》，吉川公文馆，2000年，第174页）。然而，即使到了近代明治时期，儒家道德也不允许妇女有选举权，二战战败后，“妇女解放”被总司令部列为五大维新（日本战后改革）的第一项。（第270页）被纳入其中，并且妇女被赋予了投票权。

来到现代社会，根据《2022年全球性别差距报告》公布的性别差距指数，日本在全球146个国家中排名第116位，在七国集团（G7）中遥遥垫底^[7]。

日本人普遍的性别平等意识也很薄弱。在2021年10月31日举行的第49届众议院大选中，NNN对“特别重要的政策”进行了出口民意调查，发现只有8.3%的青少年回答“促进性别平等最重要”，20多岁的人为6.5%，30多岁的人为2.5%，40多岁的人为1.5%，70岁以上的人为0.6%。^[8]

1.3 日本工作环境中的加班文化

日本的工作制度可以追溯到1930年代末以及二战期间。当时由于公司急需大量能够忍受工厂中的恶劣条件的男性员工（女性先是当秘书，然后找到丈夫后就成了家庭主妇，这通常发生在工作中），大财阀（Zaibatsu）尝试着通过用慈善的家长主义（Paternalism）对待雇员，将整个公司建成一个大家庭。作为绝对忠诚的回报，大公司的员工可以获得定期加薪、丰厚的福利和终身就业保障，因此公司关系有时比家庭关系更牢固。^[3]

第二次世界大战结束时，家庭主义进一步发展，几乎每个企业都设立了“产业报效国家组织”（Sangyo Hokokukai）该组织的口号就是“爱企如家”（Enterprise-as-a-family）。工薪族成为日本经济奇迹的冲击部队，在高速增长的时代重建了国家^[4]，却也导致了日本文化对“工作生活平衡”概念的缺失。

表一 二战前日本劳动者平均每年工作时长^[22]

1923	3252
1924	3286
1925	3231
1926	3205
1927	3168
1928	3196
1929	3169
1930	3058
1931	3037
1932	3087
1933	3147
1934	3186
1935	3191
1936	3203
1937	3219
1938	3236
1939	3224

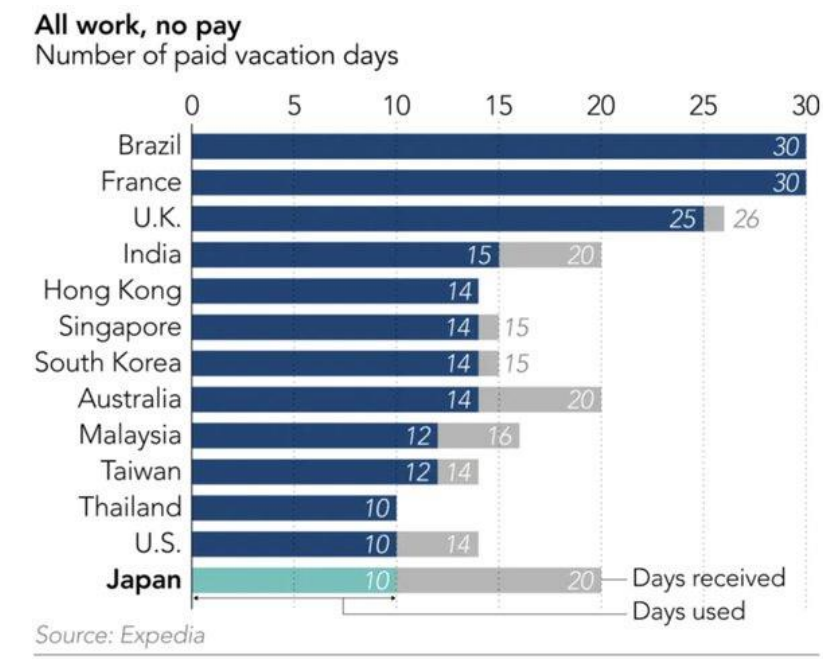
资料来源: Rodosho, Rodo Tokei Chingin Maitsuki Chosa[每月劳动调查]; Nichigin, Rodo Tokei[劳动统计], 历年。

日本工作文化的其中一个特点便是长时间工作和对公司的强烈奉献精神。这样的加班文化及其体现出的儒家思想、封建主义的等级关系，东亚集体主义思想，深深植根于日本社会。

日本的成功“工作”传统上被定义为“工作时间”，而不一定是“工作产出”。公司鼓励员工为了集体的利益而做出牺牲，称之为“坚持不懈”（gaman

suru），甚至以主动不休假期为荣。按时回家通常是懒汉在公司没有多少前途的标志，许多公司仍然普遍使用考勤卡。

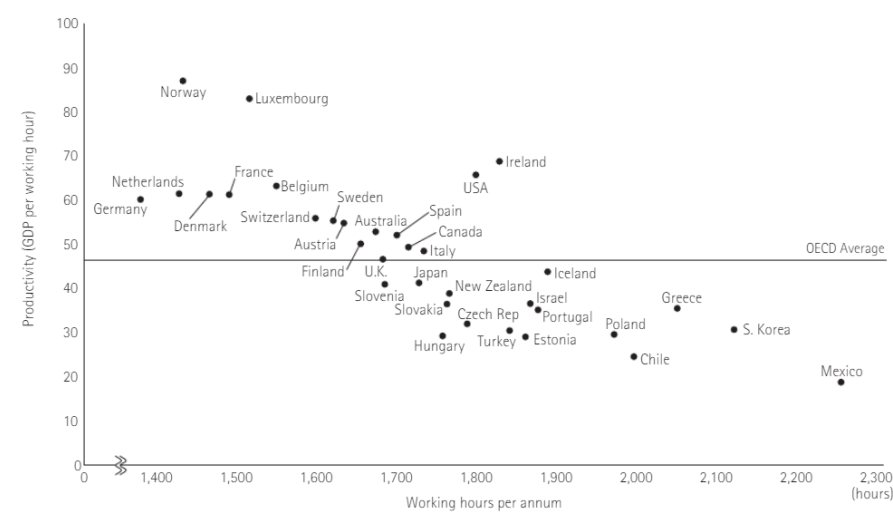
表二 部分国家/地区劳动者每年带薪假的平均使用天数^[23]



除了文化原因，造成日本加班严重的原因还有很多，已研究出的法律制度问题就有：年功序列制（一种根据工作年限和员工年龄进行工资分配的制度）、企业偏好新卒、临时工制度不发达、解雇困难等。

另外，日本还是G7中生产力和生产效率最低的国家^[8]。其产业高度以本国为中心，而不是全球性的。

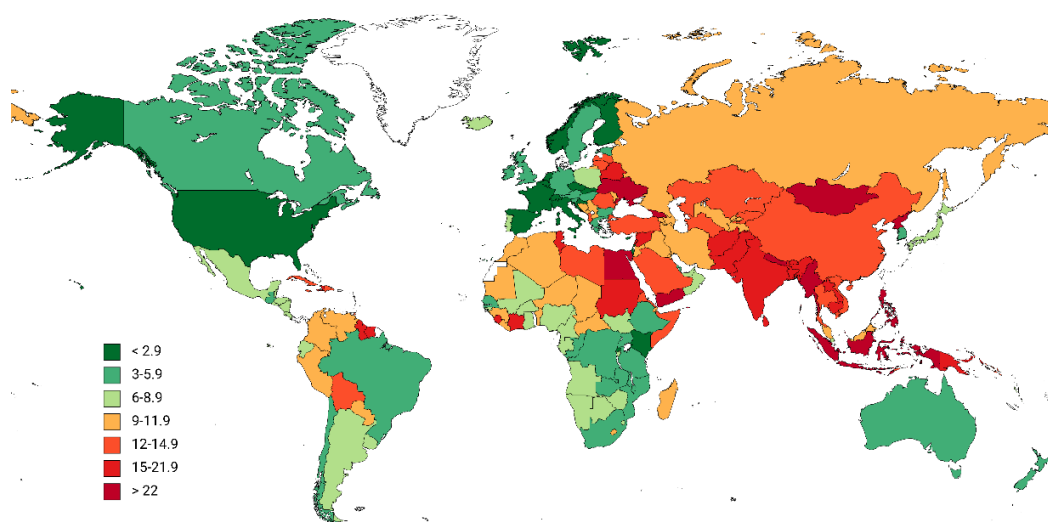
表三 OECD成员国每小时GDP产出/每年工作时长^[24]



过劳死（Karoshi，過労死），该词源自日语^[5]，是一种职业性的突然死亡，即：因过度劳动导致积劳成疾而死。

过劳死亡现象在亚洲其他地区也很普遍。根据世界卫生组织/国际劳工组织的数据，2016 年全球有 745,194 人死亡与长时间工作有关。^[6]

2015 年圣诞节，时年 24 岁的日本大型广告公司电通公司女员工高桥松里（Matsuri Takahashi）因工作过度劳累而自杀，引起了日本群众的高度愤怒，掀起了“不再过劳死”运动（No More Karaoshi），导致日本政府开启了一系列减少工作时间的改革，提高了社会对职业女性的关注度。



图二 每十万人中因工作时间过长而死亡的各国人数（15岁以上）

1.4 提出研究问题

目前没有看到使用时间序列分析知识对日本未来工作时间走势的预测。

如果以工作生活平衡情况上的两性差异为切入点进行数据分析，就能够很好地体现和预测日本劳动者总体的工作生活平衡情况，以便政府制定和调整法律政策。

二、文献综述

2.1 工作生活平衡的影响因素

为了更深入地分析，我们借用伦敦国王学院博士David Guest曾提出的用于衡量工作生活平衡质量的模型^[3]，重点关注其中的三个客观指标（objective indicators）：工作时长，自由支配时间，在家庭中扮演的角色。

表四 David Guest提出的工作生活平衡模型

<i>Determinants</i>	<i>Nature of the balance</i>	<i>Consequences/impact</i>
Organizational factors	Subjective indicators	
Demands of work	Balance – emphasis	Work satisfaction
Culture of work	equally on home and work	Life satisfaction
Demands of home	Balance – home central	Mental health/well-being
Culture of home	Balance – work central	Stress/illness
Individual factors		Behaviour/performance at work
Work orientation	Spillover and/or interference of work to home	Behaviour/performance at home
Personality	Spillover and/or interference of home to work	Impact on others at work
Energy		Impact on others at home
Personal control and coping		
Gender		
Age	Objective indicators	
Life and career stage	Hours of work	
	“Free” time	
	Family roles	

为此我们将劳动者的“性别”也纳入分析范围内，因为这与“在家庭中扮演的角色”指标密切相关。由于找不到数据，我们也忽略“年龄”“职业”对工作生活平衡的影响，即：假设所有适龄劳动者不因年龄和职业改变其工作与生活的平衡情况，在本文中不作讨论。

2.2 相关研究的综述

Mark D. West认为，日本劳动法在决定人们工作时间方面起着重要作用，特别是对于大型企业。在20世纪50年代开始，日本法院对就业市场进行了调控，使得大多数企业，尤其是大型企业，几乎不可能解雇员工。因此，大型企业雇佣了比实际需要少得多的员工，并依靠这些经过严格筛选的员工的工作时间作为缓冲。在不景气时期，员工规模使得解雇变得不必要；而在景气时期，员工被迫加班工作。^[7]

小野浩（Hiroshi Ono）认为，日本的长工作时间是日本式就业实践和文化特点的副产品。他严厉批评日本的就业制度为“过时的制度”，并重点指出，要减少工作时间，需要从输入导向转变为产出导向，实施基于绩效的奖励系统，明确工作职责并促进专业化，以及通过改变企业文化来鼓励员工减少工作时间。^[11]

Chie Aoyagi认为，日本的正式员工通常被期望遵循理想的员工形象，这往往要求他们把工作放在个人生活之前。因此，男性更有可能努力达到理想的员工形象，而女性更有可能从事非正规就业，以减少工作并照顾家庭。Chie Aoyagi还进行了一个量表实验，回归分析结果表明，虽然生产率较高的男性比生产率较低的男性更喜欢传统的任务分工，但生产率较高的女性比生产率较低的女性更不喜欢传统的选择。这可能提醒我们，劳动力市场改革的成功可能至少在一定程度上取决于改革方案是否解决了性别刻板印象。^[15]

Maiko Hata将日本准备转变为“女性经济学”的官方声明与日本公司的实际成就和面临的挑战进行对比，认为公司将陪产假作为其努力的一部分，能创建一个对潜在员工更具吸引力的工作场所，以及一个可以以“多元化”或“平等就业机会”雇主的标签为荣的工作场所。^[17]

Kuniko Ishiguro回顾了 1980 年代中期以来日本职场性别平等的发展，认为平等就业机会法在改变社会和经济对性别平等的看法方面发挥了重要作用。然而，公司在工资水平和其他领域仍然存在很大的性别差距。^[14]

Rie S. Hori统计后发现，日本 STEM 领域女性比例改善缓慢且偏低，**尤其是在数学领域**。2018 年获得数学博士学位的女性比例仅有 6%（150 人中有 9 人），是过去 20 年来最低的一年。作为对比，韩国在2006年的比例为42.1%。^[16]

2.3 当前研究的创新点

本文尝试运用数据分析的相关知识，着重研究以不同的公司规模下日本劳动者的工作时长特征，旨在东亚发达国家视角下，构建出关于工作与生活的类像。

2.4 数据的来源和介绍

本报告中的数据取自国际劳工组织（International Labour Organization，简称ILO）的统计数据库ILOSTAT和日本政府统计门户网站e-stat，编号00200531，2-6-9，2-8-8。表格标题为 Employed person by type of household, relationship to the head of household, weekly hours of work and wishing to increase/decrease worktime(Since Jan. -Mar. 2011) 以及 Average/Aggregated weekly hours of work at by duration of employment contract (Since January 2018)，编号0003024371，最终得到共 6687 份 2024年1月到3月的15岁以上适龄日本劳动者数据，以及共 274 个月份的男女月均工作时间数据。

表格内容包含“性别”“学历”“家庭角色”“每周平均工作时间”“是否希望增加/减少工作时间”，属于计数资料。

对就业的定义为：不包括自用生产工人。

工作范围：目前从事的主要工作。

家庭角色：「戸籍謄本」与中国的户口本类似，但也有不同。戸籍謄本是以一对夫妇为单位设立的，孩子出生后，名字会被登记在同一户籍中。孩子结婚后就算仍然还住在同一个家中，也必须从原有

的户籍中脱离出来，重新建立一个新的户籍。

第三章主要研究于2024-05-14 14:00发布的最新数据，统计时间为日本2024年春季（1月到3月）。第四章则是对十年间的月均工作时间进行了时间序列分析。

由于数据库中没有数据的具体分布情况，不涉及总体的参数（均值、方差），因此对现状的分析将使用非参数检验，然后会以时间为自变量进行参数检验。

三、性别和家庭角色的非参数检验

3.1. 描述性统计分析

表五 样本人口学汇总表（n=6687）			
人口学变量		频数（万人）	百分比
性别	男	3657	54.69%
	女	3030	45.31%
学历	毕业	6340	94.81%
	未毕业	347	5.19%
年龄	15 到 24 岁	529	9.17%
	25 到 34 岁	1110	19.25%
	35 到 44 岁	1266	21.96%
	45 到 54 岁	1613	27.98%
	55 到 64 岁	1248	21.64%
家庭角色 （与户主的关系）	两人及以上家庭	5598	45.57%
	户主	2489	20.26%
	户主的配偶	1667	13.57%
	孩子或孩子的配偶	1315	10.70%
	其他亲戚	126	1.03%
	一人家庭	1089	8.87%

标准差，均值等未知。

3.2 相关性分析

3.2.1 分析概述

由于原始数据中无连续型变量，因此我们无法列出相关系数矩阵。但我们

仍然可以通过非参数检验进行间接的相关性分析。改变工时意愿属于等级资料，其中：等级1表示希望增加工作时间，等级2表示不希望工作时间改变，等级3表示希望减少工作时间。

3.2.2 卡方检验

由于我们只有频数表，而没有调查问卷的原始数据，因此我们无法确定样本标准差，也无法确定每周平均工作时间是否服从正态分布，因此无法使用t检验等参数检验方法。但我们仍然可以使用卡方独立性检验等非参数检验方法。

卡方检验适用于频数数据，即分类数据的观测值。显然不同样本的周均工时满足独立性，且样本量和期望频数 T 足够大 ($T \geq 5$)，但“是否希望增加/减少工作时间”则不一定满足。下面以性别与每周平均工作时间为例，过程如下：

首先提出假设：

原假设(H0)：假设两个变量之间没有关联，即它们是独立的。

备择假设(H1)：假设两个变量之间存在关联，即它们不是独立的。

然后整理表格，收集数据，构建列联表，记录两个分类变量的频数。

根据行和列的边际总和计算期望计数。公式为：

$$E_{ij} = \frac{(R_i \times C_j)}{N}$$

其中 E_{ij} 是第*i*行第*j*列的期望计数， R_i 是第*i*行的总和， C_j 是第*j*列的总和， N 是总样本数。

在SPSS软件输入表格，对人数进行加权后进行交叉制表，得到：

表六 周均工时 * 性别 交叉表					
		性别		总计	
		男	女		
周均工时	0-34小时	计数	961	1636	2597
		期望计数	1420.3	1176.7	2597.0
		占 周均工时的百分比	37.0%	63.0%	100.0%
	35小时以上	计数	2696	1394	4090
		期望计数	2236.7	1853.3	4090.0
		占 周均工时的百分比	65.9%	34.1%	100.0%
总计		计数	3657	3030	6687
		期望计数	3657.0	3030.0	6687.0
		占 周均工时的百分比	54.7%	45.3%	100.0%

然后利用spss或者公式计算卡方统计量，公式为：

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

其中 O_{ij} 是第 i 行第 j 列的观察频数， E_{ij} 是第 i 行第 j 列的期望频数。

并且确定自由度：

$$df = (r - 1) \times (c - 1)$$

其中 r 是行数， c 是列数。

得到表格：

表七 SPSS输出的卡方检验结果样例

	值	自由度	渐进显著性 (双侧)	精确显著性 (双侧)	精确显著性 (单侧)
皮尔逊卡方	535.839 ^a	1	.000		
连续性修正 ^b	534.673	1	.000		
似然比(L)	540.373	1	.000		
费希尔精确检验				.000	.000
线性关联	535.759	1	.000		
有效个案数	6687				

a. 0 个单元格 (0.0%) 的期望计数小于 5。最小期望计数为 1176.75。

b. 仅针对 2x2 表进行计算

由于 $N \geq 40$ ， $T \geq 5$ ，我们使用Pearson卡方检验结果。比较计算出的 p 值与显著性水平（如 0.05）。如果 p 值小于显著性水平，则拒绝原假设（即认为两个分类变量之间存在显著的关联）。

计算出的 p 值非常小，所以我们拒绝原假设，表示在原假设成立的情况下，观察到的检验统计量值或更极端的值的概率很小，有足够的证据认为性别和每周平均工作时间之间存在显著性关联。

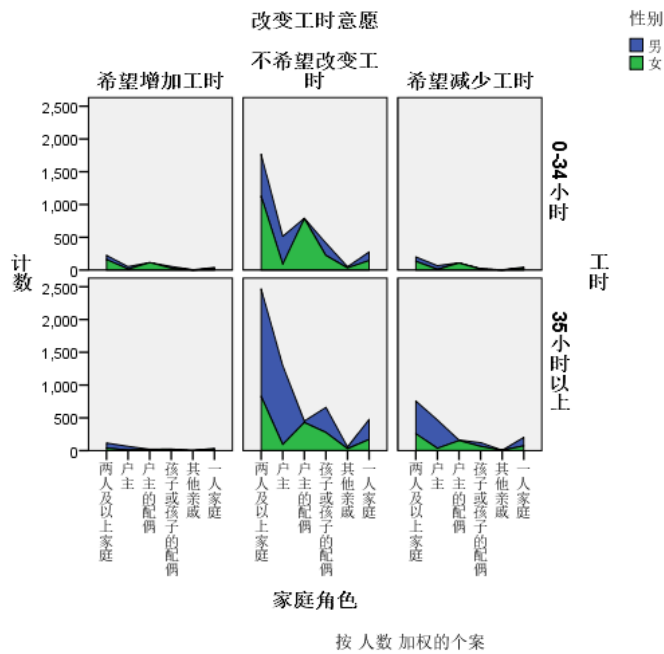
将四个变量两两列联成交叉表，分别进行卡方检验，将得到的P值整合成一个四维矩阵，可以初步肯定，四个变量之间均具有强相关性。

表八 P值矩阵表

p	性别	家庭角色	工时	改变工时意愿
性别				
家庭角色	0.000			
工时	0.000	0.000		
改变工时意愿	0.000	0.000	0.000	

3.3 可视化分析

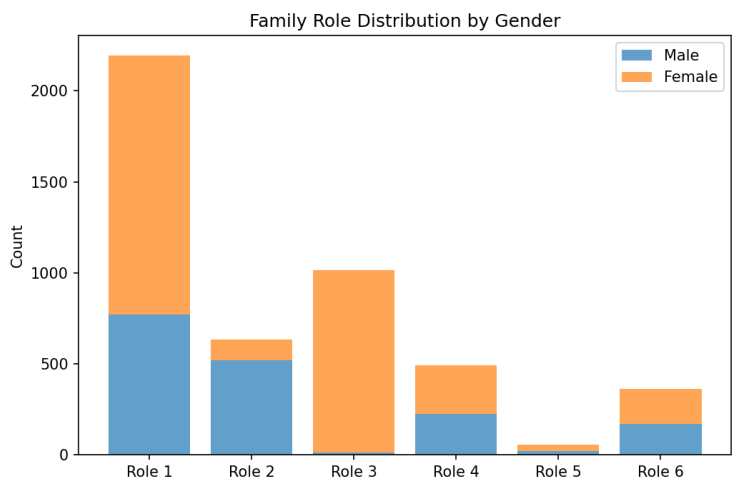
图三 样本总频数堆叠图



可以看到，角色一（两人或两人以上的家庭）中的女性比例异常高于男性，说明了当今日本社会女性与男性共同工作支撑家庭的趋势。男性更多承担角色二（户主）的位置，而角色三（户主的配偶）则几乎全部由女性承担吗，体现了日本保守社会倾向于男性来扮演一家之主的角色。

在角色四（孩子及孩子的配偶）、角色五（其他亲戚）、角色六（独身家庭）中，男女比例近似一致，符合我们对数据的期望，验证了数据的可信度。

图四 以性别划分的家庭角色条形图



3.4 性别差异分析与假设检验

对性别与每周平均工作时间进行卡方检验，P值极小。

可以拒绝原假设，认为二者来自不同的分布。

表九 性别与每周平均工作时间 列联表

周均工时	0~34h	35h+	高工时率
男	961	2696	73.72%
女	1636	1394	46.01%
χ^2			535.839 ^a
<i>p</i>			0.000

由于性别是2个独立样本，意愿为等级资料，进行Wilcoxon Rank秩和检验。

同样P值极小，因此我们拒绝原假设，认为二者来自不同的分布。

（两组的中位数不等）

表十 性别对改变工作意愿的Wilcoxon Rank秩和检验^a

改变工时意愿	
曼-惠特尼 U	17260906.000
威尔科克森 W	32615017.000
Z	-7.197
渐近显著性（双尾）	.000

a. 分组变量：性别

		秩		
	性别	N	秩平均值	秩的总和
改变工时意愿	男	6603	6228.90	41129423.00
	女	5541	5886.12	32615017.00
	总计	12144		

相较于男性，女性的秩平均值更低，意味着女性更加偏向选择增加自己的工作时长，而男性偏向于减少。

3.5 家庭角色差异分析与假设检验

由于家庭角色是6个独立样本，进行Kruskal-Wallis秩和检验。

同样P值极小，因此我们拒绝原假设，认为二者来自不同的分布。

表十一 家庭角色对改变工作意愿的Kruskal-Wallis秩和检验

改变工时意愿	
克鲁斯卡尔-沃利斯 H	81.938
自由度	5
渐近显著性	.000

a. 克鲁斯卡尔-沃利斯检验

b. 分组变量：家庭角色

秩			
	家庭角色	N	秩平均值
改变工时意愿	两人及以上家庭	5533	6046.37
	户主	2466	6362.26
	户主的配偶	1648	5889.12
	孩子或孩子的配偶	1298	5696.56
	其他亲戚	124	5545.53
	一人家庭	1075	6338.13
	总计	12144	

6个家庭角色中，户主秩平均值最高。这是因为户主承担着全家收入来源的职责，平衡工作与生活的愿望也最强烈。而配偶、孩童和亲戚在家庭中则无需承担过多的职责，因此秩平均值最低。

此外，由于一人家庭需要独自兼顾工作与家务，因此减少工作时间的愿望也很强。

3.6 工时与改变工时意愿的非参数统计检验

斯皮尔曼等级相关系数（Spearman's Rank Correlation Coefficient）是一种非参数统计方法，用于测量两个变量之间的单调关系（无论是线性还是非线性）。它的计算基于数据排名，而不是实际值，因此特别适合于处理非线性关系和不满足正态分布的数据。把工时划分为1和2，（0-34h和35h以上）

计算得到：

表十二 Spearman检验	
r_s	p
0.20923	0.00000

可以看到，尽管工时与改变工时意愿之间的相关性较弱，为0.20923，但由于p值非常低，说明这种弱相关性是显著的。p值为0.00000意味着我们有极强的证据表明相关性不是由随机噪音引起的，而是存在实际的关联。

四、月平均工作时长的时间序列分析

4.1 数据的预处理

4.1.1 描述性统计

整理2013-2024十年间的男女月均工作时间数据，并进行描述性统计得到：

表十三 2013-2024男女月均工作时间描述性统计

月工作小时								
	个案数	平均值	标准差	标准误差	平均值的 95% 置信区间		最小值	最大值
					下限	上限		
男	137	175.28905	7.80635	.66694	173.97013	176.60797	153.70000	191.40000
女	137	133.15547	4.94680	.42263	132.31969	133.99126	117.90000	143.80000
总计	274	154.22226	22.09036	1.33453	151.59499	156.84954	117.90000	191.40000

经检查无缺失值，定义时间序列为2013年1月到2024年5月。

4.1.2 正态性检验

根据中心极限定理，在许多情况下，对于独立并同样分布的随机变量，即使原始变量本身不是正态分布，标准化样本均值的抽样分布也趋向于标准正态分布。那么对日本十年间的每月工作时间进行分析，结果有很大概率是服从正态分布的。

柯尔莫哥洛夫-斯米尔诺夫检验（英语：Kolmogorov-Smirnov test，简称K-S test），是一种基于累计分布函数的非参数检验，用以检验两个经验分布是否不同或一个经验分布与另一个理想分布是否不同。

假设男女每月工作时间之间是独立的，由于：

当数据量≤50时，倾向于以S-W(Shapiro–Wilk)检验结果为准。

当50<样本量<5000时，正态性检验倾向于以K-S检验结果为准。

而样本量为137，接近50，保险起见分别对男、女进行单样本 S-W和K-S 检验。

提出原假设 H0：样本来自正态分布。

备择假设 H1：样本不来自正态分布。

表十四 对月均工作时间的正态性检验

	柯尔莫戈洛夫-斯米尔诺夫 ^a			夏皮洛-威尔克		
	统计	自由度	显著性	统计	自由度	显著性
男	.040	137	.200 [*]	.990	137	.459
女	.053	137	.200 [*]	.991	137	.496

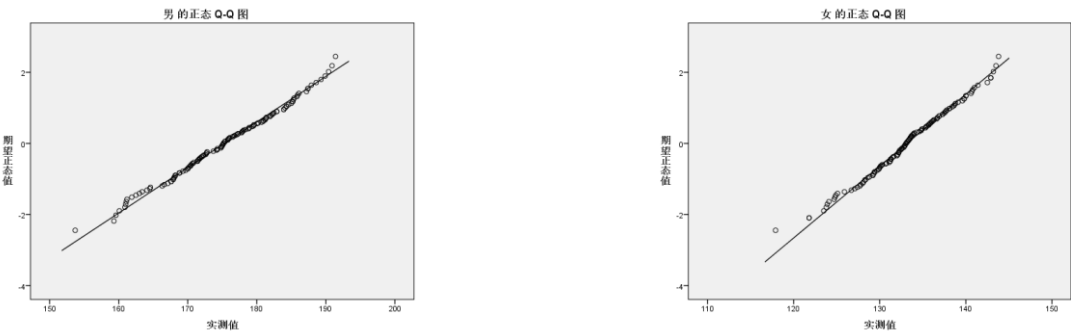
由于两个检验的P值均大于0.1，且K-S统计量较小，表明样本数据与理论分布之间没有显著差异，因此不能拒绝原假设，认为非常服从正态分布。

表十五 对月均工作时间的正态性检验结果

性别	K-S 统计量	p
男	0.04	0.200
女	0.053	0.200

观察Q-Q图，可以看到散点近似贴合正态值直线，验证了结论。

图五 月均工作时间的正态Q-Q图



4.1.3 方差齐性检验

由于标准差差异很大，男女两组的方差差异显然会是显著的。

表十六 对月均工作时间的Levene方差齐性检验

ANOVA					
月工作小时					
	平方和	自由度	均方	F	显著性
组间	121603.822	1	121603.822	2847.533	<.001
组内	11615.752	272	42.705		
总计	133219.574	273			

方差齐性检验					
		莱文统计	自由度 1	自由度 2	显著性
月工作小时	基于平均值	26.560	1	272	<.001
	基于中位数	26.506	1	272	<.001
	基于中位数并具有调整后自由度	26.506	1	239.347	<.001
	基于剪除后平均值	26.583	1	272	<.001

经Levene方差齐性检验后发现P值确实非常小，符合我们的已有结论。

4.1.4 t检验

独立样本 t 检验常用于比较两个独立样本的均值是否相等。

提出原假设 H_0 ：两个样本的均值相等。

备择假设 H_1 ：两个样本的均值不相等。

计算得到：

表十七 对月均工作时间的t检验		
t	p	效应量
53.362	4.92×10^{-7}	8.517

P值远远小于0.05，因此拒绝原假设，说明两个样本的均值显著不同。

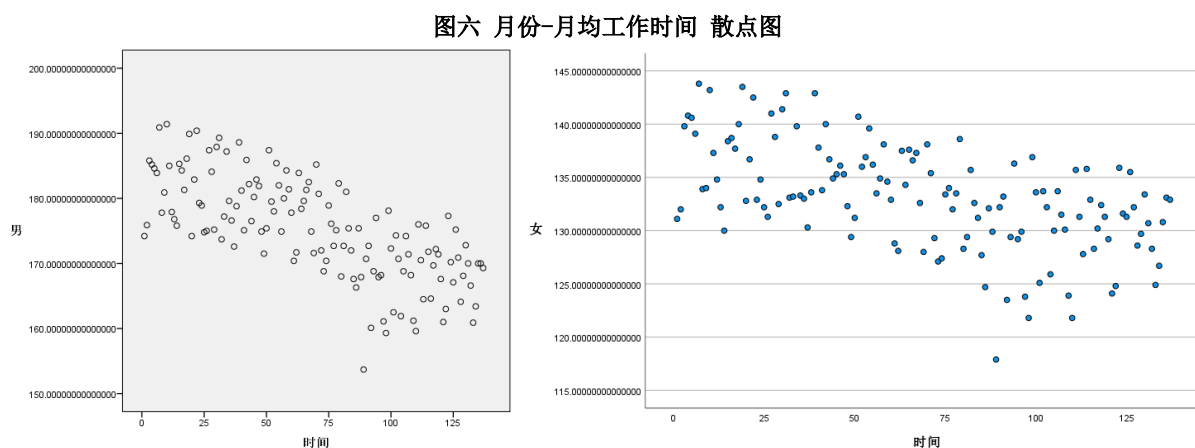
由于方差不齐，效应量选用Glass's Delta。

4.2 时间序列的回归分析

4.2.1 假设检验

经由散点图显示，猜想月均工作时间与月份服从直线线性关系。下面以对男性的分析为例介绍了回归分析的过程。

在SPSS软件绘制月份与月均工作时间的散点图，初步猜想二者具有线性关系，尝试构建时间序列的回归模型。



提出假设 H_0 :回归方程无统计学意义， $\alpha=0.05$ 。计算得出 $F=132.835, p<0.001$,

因此拒绝 H_0 ,回归方程有统计学意义。得到方程 $Y=184.846-1.39x$ 。

同样步骤得出女性的回归方程为 $Y=137.947-0.069x$ 。

4.2.2 系数检验

提出假设 $H_0: \beta = 0$ ，无直线回归关系， $\alpha=0.05$ 。计算得出 $T=-11.525, p<0.001$,

因此拒绝H0,认为有直线回归关系。

4.2.3 残差分析

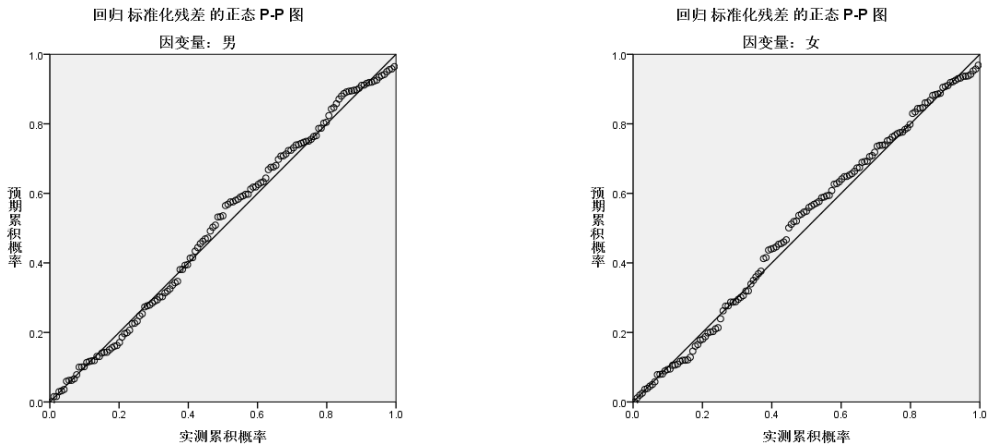
对残差进行统计得到标准偏差为0.996，在月均工作时长（100到200）的区间范围内残差值较小，认为模型拟合程度尚可。

表十七 对线性回归模型（男）的残差统计^a

	最小值	最大值	平均值	标准偏差	个案数
预测值	165.871	184.707	175.289	5.498	137.000
残差	-18.819	10.049	.000	5.542	137.000
标准预测值	-1.713	1.713	.000	1.000	137.000
标准残差	-3.383	1.807	.000	.996	137.000

a. 因变量：男

图八 回归方程的标准化残差统计 正态P-P图

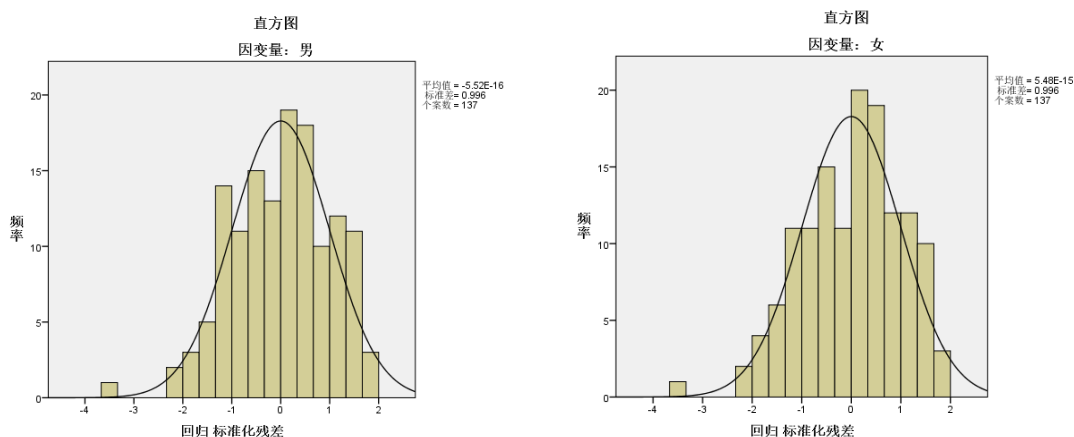


观察正态P-P图，可以看到点沿对角线浮动，残差近似服从正态分布，且分布没有明显的偏差。这表示样本的标准化残差和正态分布吻合得较好。

德宾-沃森（Durbin-Watson）统计量是一种用于检验回归模型中残差是否存在自相关的统计检验方法。它的取值范围通常在0到4之间，其值越接近2表示残差之间的自相关性越小（接近0表示存在正自相关，接近4表示存在负自相关）

对残差进行Durbin-Watson检验，得到D-W统计量为1.705，接近于2，表明残差之间没有或只有很小的自相关。

图七 回归方程的标准化残差统计 直方图



正态直方图再次验证了数据的正态性。

可能的缺陷是，浮动虽然很小，但并不是均匀分布在对角线附近，呈现一定规律的振荡。说明残差仍有一定可能不是白噪声，存在某种系统性偏差。

4.2.4 结果解读

总结得到：

表十八 线性回归方程的拟合结果

性别	F	P	回归方程	置信区间	T	决定系数 R^2
男	132.835	0.000	$Y=184.846-1.39x$	$(-0.162, -0.115)$	-11.525	0.496
女	60.683	0.000	$Y=137.947-0.069x$	$(-0.087, -0.052)$	-7.798	0.311

可以看出，2013年时男性的工作时间显著高于女性，平均每月比女性多工作47小时，但经过十年的改革，相比女性，工作时长正在持续地锐减，减少速率是女性的20倍。

而女性的工作时间虽然一直较少，但在十年间的减少速度不大，平均仅减少了0.69小时，这也许表示在女性在职场中越来越多地开始扮演中坚角色，而过去职场鼓励女性成为家庭主妇的现象在减少，是日本性别平权进步的体现。

4.3 基于ARIMA模型的时间序列分析

4.3.1 ARIMA模型概况

平稳的时间序列的性质不随观测时间的变化而变化。^[20]因此具有趋势或季节性的时间序列不是平稳时间序列——趋势和季节性使得时间序列在不同时段呈现不同性质。与它们相反，白噪声序列（white noise series）则是平稳的——不管观测的时间如何变化，它看起来都应该是一样的。

自回归移动平均模型(AutoregressiveIntegrated Moving AverageModel, ARIMA)是由 Box和 Jenkins在20世纪70年代初提出的时间序列检测方法，又称为BoxJenkins模型

^{[13][21]}。它在统计学和计量经济学等多种学科当中都有所应用,是最常见的一种用于进行时间序列预测的方法。^[21]

ARIMA 模型是几种情况的综合表述形式,其具体取决于 3 个方面:是否由时序数据本身的回归构成;是否由随机误差项的回归构成;是否对序列做了差分。时序数据在平稳性的基础上才能对具体模型形式进行判别。^[18]因此 ARIMA 模型 根据实际情况具体表现为:滑动平均模型(MA)、自回归模型 (AR)、回归G 滑动平均模型(ARMA)、差分后的滑动平均自 回归模型(ARIMA)。

其中:

- AR部分用于处理时间序列的自回归部分,它考虑了过去若干时期的观测值对当前值的影响。
- I部分用于使非平稳时间序列达到平稳,通过一阶或者二阶等差分处理,消除了时间序列中的趋势和季节性因素。
- MA部分用于处理时间序列的移动平均部分,它考虑了过去的预测误差对当前值的影响。

形式上看,ARIMA模型的公式可以表示为:

$$Y_t = c + \varphi_1 Y_{t-1} + \varphi_2 Y_{t-2} + \dots + \varphi_p Y_{t-p} + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \dots + \theta_q \epsilon_{t-q} + \epsilon_t$$

在这个公式中:

Y_t 是我们正在考虑的时间序列数据。

φ_1 到 φ_p 是AR模型的参数,用来描述当前值与过去p个时间点值之间的关系。

θ_1 到 θ_q 是MA模型的参数,用来描述当前值与过去q个时间点的误差之间的关系。

ϵ_t 是在t时间点的误差项。

c 是一个常数项。

简而言之,ARIMA模型就是试图通过数据的自相关性和差分的方式,提取出隐藏在数据背后的时间序列模式,然后用这些模式来预测未来的数据。

4.3.2 实验环境

本模型所有运行步骤均在PC上执行,SPSS软件版本为27.0,PC配置环境如下:

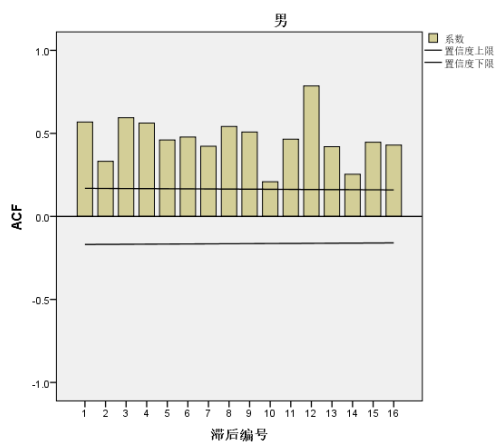
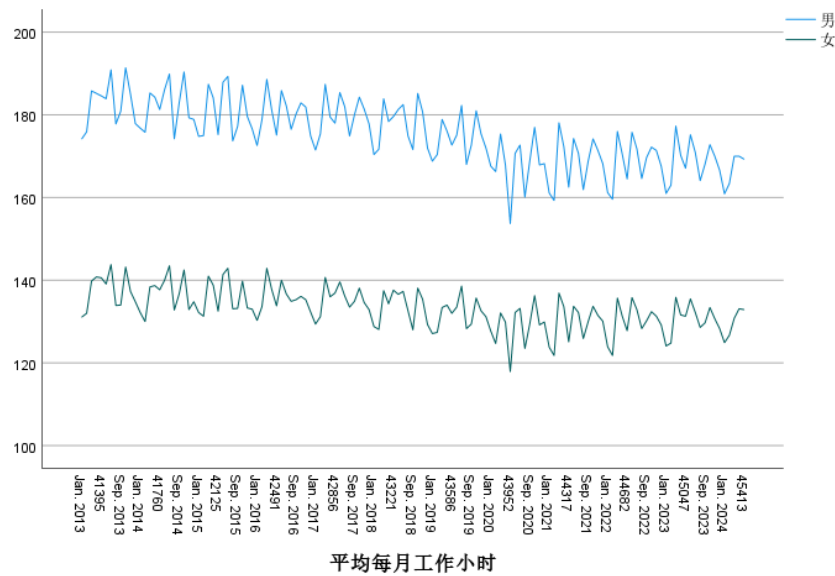
表十九 PC配置环境

组件	规格
系统	Windows 11 家庭版 64 位 版本号 22H2
处理器	Intel Core i5-12500H 核心数: 12 (4P+8E), 线程数: 16 (8P+8E), 工艺: Intel 7
显卡	Intel Iris Xe Graphics NVIDIA GeForce RTX 2050 Lenovo 流处理器: 640, 显存: 1GB 流处理器: 2048, 显存: 4GB (GDDR6 Samsung)
主板	LNVNB161216 LENOVO, 芯片组: Intel Alder Lake-P PCH
硬盘	Micron MTFDKBA512TFH, 实际容量: 477GB, 类型: SSD
显示器	Lenovo LEN160WQ, 分辨率: 2560*1600, 刷新率: 60Hz, 屏幕尺寸: 15.9 英寸
内存	Samsung DDR5-6400 (3200 MHz) 2GB x 8 条 容量: 16GB, 通道: 4, 频率: 4788 MHz,

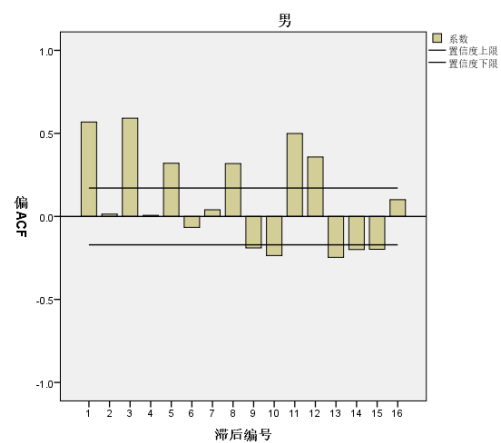
4.3.3 平稳性检验

对男女每月工作时间进行单样本检验，经检查无缺失值，定义时间序列为2013年1月到2024年5月，创建时间序列，得到原始时间序列图，ACF图和PACF图。

图八 原始时间序列图



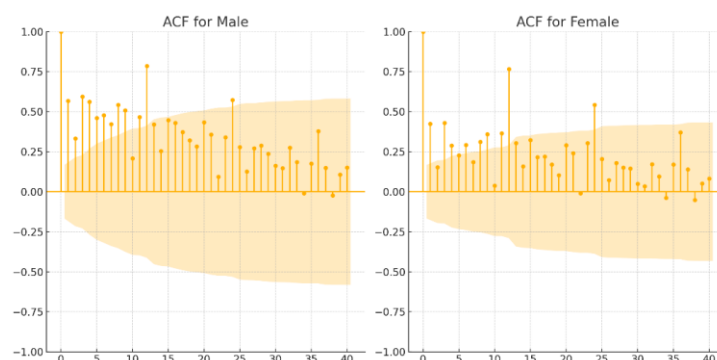
图九 原始时间序列ACF图（男）



图十 原始时间序列PACF图（男）

由图可知，ACF是拖尾，没有向零收敛的趋势，而是在置信度上限之外不断波动；PACF为1阶截尾，系数围绕零轴上下大跨度地波动。

图十一 原始时间序列ACF与显著性区间图



图十一中，黄色区域表示显著性区间，超过该区域的自相关系数表示显著自相关。可以看到ACF值在多个滞后期内显著，说明男女时间序列原始数据都有显著的自相关性，表明数据具有趋势或周期性。

因此判定原始时间序列为非平稳序列，需要先进行d阶差分将其转化为平稳序列。为确定d值，分别对男、女进行 ADF检验（单位根检验）。

表二十 ADF检验表

变量	差分阶数	t	p	AIC	临界值		
					1%	5%	10%
男	0	-0.6965	0.8476	630.52	-3.4847	-2.8853	-2.5795
男	1	-3.3025	0.0148	620.8	-3.4851	-2.8855	-2.5796
男	2	-10.1165	9.69E-18	625.24	-3.4851	-2.8855	-2.5796
女	0	-1.2381	0.657	571.55	-3.4838	-2.885	-2.5793
女	1	-13.6655	1.49E-25	567.27	-3.4838	-2.885	-2.5793
女	2	-10.8377	1.65E-19	581.17	-3.4851	-2.8855	-2.5796

对于“男”变量，差分阶数为2时t统计量显著（p值远小于0.05），说明在二次差分后时间序列平稳。

对于“女”变量，差分阶数为1和2时t统计量显著（p值远小于0.05），说明在一次差分和二次差分后时间序列平稳。

但，男女两变量的AIC（Akaike信息准则）最低值均在一阶差分处，而AIC值越小通常会代表模型越好。

因此，尽管0.0148并没有远远小于0.05，但我们仍然可以拒绝单位根的存在假设，通过AIC值综合判断d值为1。

4.3.4 白噪声检验

计算一阶差分的后的序列ACF和PACF值，并观察ACF图表，发现自相关性确实已经较好，没有显著超出阈值的峰值，认为是该序列为平稳非白噪声序列，具有分析价值。

表二十 一阶差分后的时间序列（男）自相关与偏自相关检验表

自相关性						偏自相关性		
序列: 男						序列: 男		
延迟	自相关性	标准误差 ^a	博克斯-杨统计			延迟	偏自相关性	标准误差
			值	自由度	显著性 ^b			
1	-.227	.085	7.161	1	.007	1	.297	.089
2	-.579	.084	54.090	2	<.001	2	.149	.089
3	.338	.084	70.177	3	<.001	3	.329	.089
4	.078	.084	71.042	4	<.001	4	-.170	.089
5	-.134	.084	73.612	5	<.001	5	.039	.089
6	.088	.083	74.729	6	<.001	6	.236	.089
7	-.202	.083	80.663	7	<.001	7	.025	.089
8	.174	.083	85.116	8	<.001	8	.037	.089
9	.307	.082	99.032	9	<.001	9	-.056	.089
10	-.642	.082	160.422	10	<.001	10	-.189	.089
11	-.072	.082	161.199	11	<.001	11	.060	.089
12	.792	.081	256.115	12	<.001	12	-.171	.089
13	-.230	.081	264.207	13	<.001	13	.136	.089
14	-.411	.081	290.141	14	<.001	14	.178	.089
15	.232	.080	298.526	15	<.001	15	-.055	.089
16	.045	.080	298.837	16	<.001	16	-.042	.089

a. 假定的基本过程为独立性（白噪声）。

b. 基于渐近卡方近似值。

4.3.5 季节性分析

对序列进行季节性分解技术STL分解（Seasonal and Trend decomposition using Loess），通过局部加权回归（Loess）将时间序列分解为趋势、季节性和残差三个部分。得到各月份的季节因子，季节因子是根据不同月份对于工作时间的影

表二十一 原始序列的季节因子表

季节因子		
序列名称	周期	季节因子 (%)
男	1	95.8
	2	96.5
	3	103.7
	4	100.7
	5	98.0
	6	102.8
	7	102.7
	8	96.8
	9	99.7
	10	103.6
	11	101.0
	12	98.7
女	1	95.9
	2	95.6
	3	103.3
	4	101.1
	5	99.1
	6	103.0
	7	103.0
	8	98.0
	9	99.2
	10	103.0
	11	100.2
	12	98.8

在上文中，我们知道，即使在去除季节性影响，没有进行季节性差分的情况下，数据在给定的显著性水平下仍然是平稳的，这间接表明季节性成分不显著，与我们的STL分析结果矛盾。究竟谁是对的呢？答案将在模型检验中给出。

4.3.6 模型估计

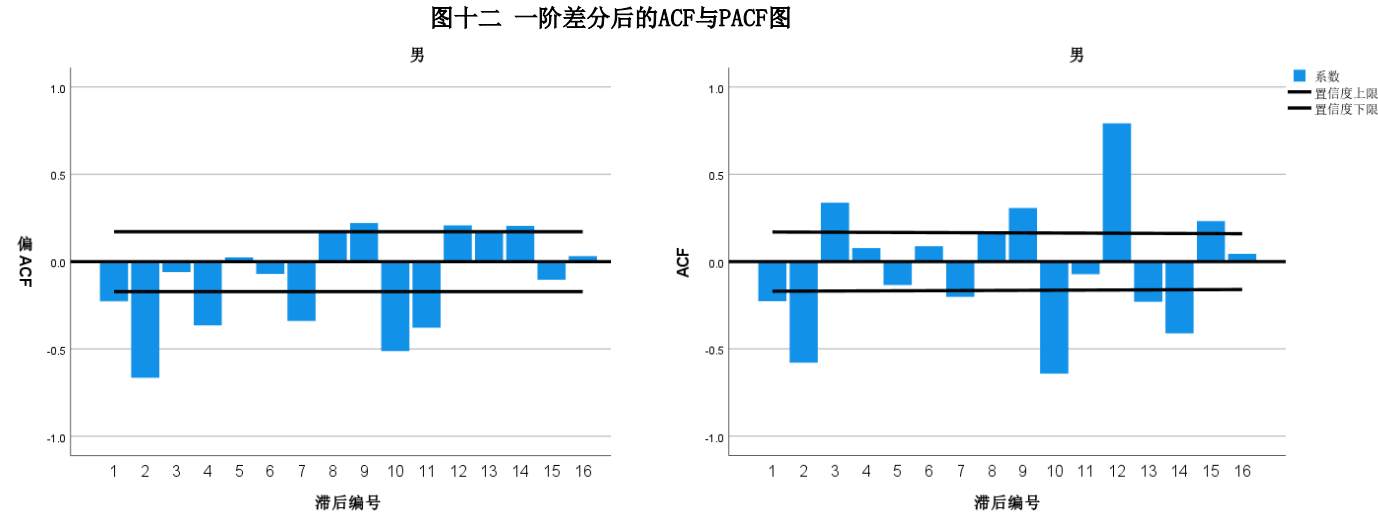
移动平均阶数 q 指的是模型中包含的前几期的移动平均项。在确定 q 值时，主要依据是时间序列的自相关函数（ACF）图。

观察ACF图，在显著水平下第一次超出阈值的滞后项为2，选择 q 值为2。

自回归阶数 p 指的是模型中包含的前几期的自回归项。在确定 p 值时，主要依据是时间序列的偏自相关函数（PACF）图。

观察PACF图，在显著水平下第一次超出阈值的滞后项为2，选择 p 值为2。

综上所述，我最终选择了ARIMA模型（2,1,2）以及SARIMA模型（2,1,2）进行拟合。



4.3.7 模型拟合

4.3.7.1 ARIMA模型

在SPSS上进行拟合，得到ARIMA模型数据。

表二十一 ARIMA模型概要（详见附录）

模型统计							
模型	预测变量数	模型拟合度统计		杨-博克斯 Q(18)			
		平稳 R 方	R 方	统计	DF	显著性	离群值数
男-模型_1	1	.580	.639	69.338	14	<.001	0
女-模型_2	1	.523	.451	71.336	14	<.001	0

ARIMA 模型参数								
				估算	标准误差	t	显著性	
男-模型_1	男	不转换	常量		-.113	.103	-1.095	.276
			AR	延迟 1	-.369	.105	-3.504	<.001
				延迟 2	-.450	.092	-4.904	<.001
			差异	1				
			MA	延迟 1	.227	.103	2.202	.029
				延迟 2	.578	.102	5.654	<.001
女-模型_2	时间	不转换	分子	延迟 0	-5.599E-5	.001	-.042	.966
	女	不转换	常量		-.106	.037	-2.843	.005
			AR	延迟 1	-.436	.108	-4.044	<.001
				延迟 2	-.240	.098	-2.458	.015
			差异	1				
			MA	延迟 1	.192	3.496	.055	.956
				延迟 2	.807	2.834	.285	.776
	时间	不转换	分子	延迟 0	.001	.001	.906	.367

Q统计量是对残差序列进行一阶至某阶自相关检验的统计量。在时间序列分析中，如果残差序列是白噪声（即无自相关性），则Q统计的值应该在一定的显著性水平下服从卡方分布。

t统计量则是用来评估每个自回归系数是否显著不为零。

在模型统计中，可以看到，ARIMA模型的拟合优度（ R^2 ）较高，达到0.639和0.451，说明拟合程度较好，但Q统计量的显著性很低，说明残差序列存在显著的自相关性，即残差并不符合白噪声的特性。

在模型参数中，可以看到，AR模型对男性和女性的拟合度很好，p值均小于0.001，显著性强；MA模型对男性的拟合度也很好，但对女性拟合效果一般。

4.3.7.2 SARIMA模型

在SPSS上进行拟合，得到SARIMA模型数据。

在模型统计中，可以看到，SARIMA模型的拟合优度（ R^2 ）更高，能达到0.837和0.794，说明拟合程度比ARIMA模型好，且女性Q统计量的显著性比ARIMA略高。

在模型参数中，可以看到，AR模型对男性拟合度一般，女性的拟合度非常好；MA模型对男性的拟合度也比较一般，对女性拟合度很好。

表二十二 SARIMA模型概要（详见附录）

模型统计							
模型	预测变量数	模型拟合度统计		杨-博克斯 Q(18)			离群值数
		平稳 R 方	R 方	统计	DF	显著性	
男-模型_1	1	.244	.837	61.838	14	<.001	0
女-模型_2	1	.325	.794	25.930	14	.026	0

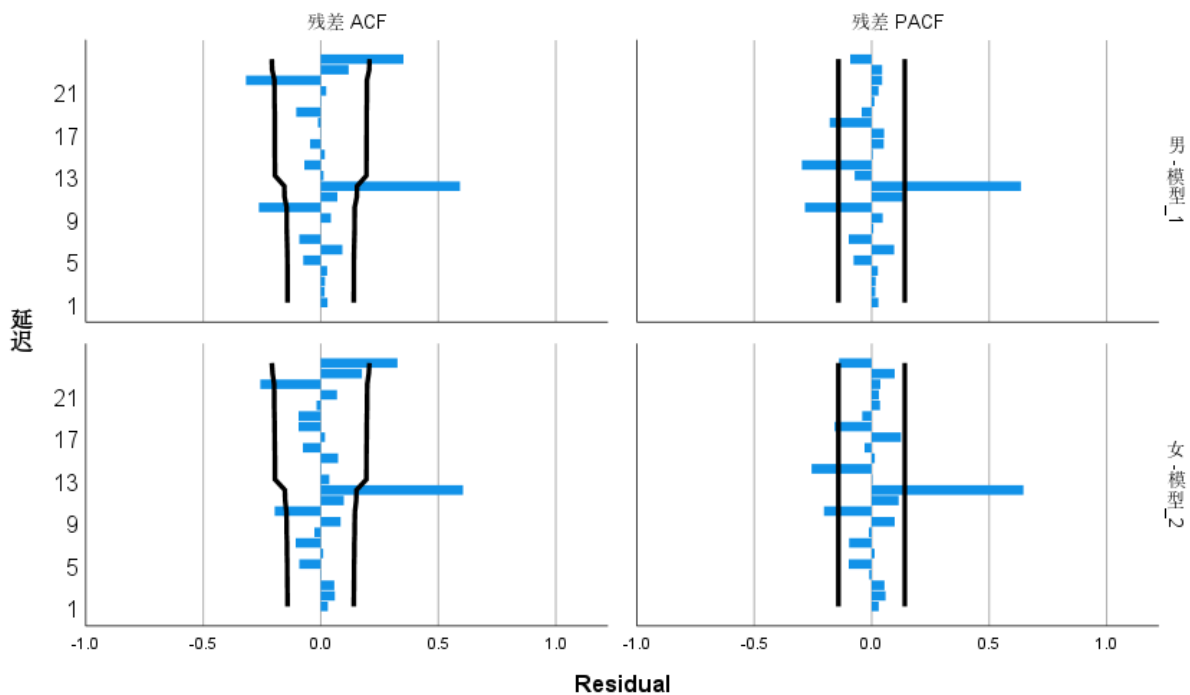
ARIMA 模型参数

					估算	标准误差	t	显著性
男-模型_1	男	不转换	常量		-1.558	.455	-3.425	<.001
			AR, 季节性	延迟 1	.349	.336	1.042	.300
				延迟 2	-.213	.275	-.774	.440
			季节性差异		1			
			MA, 季节性	延迟 1	.775	1.468	.528	.598
				延迟 2	.199	.650	.306	.760
			时间	不转换	分子			
			延迟 0		-.001	.006	-.189	.850
女-模型_2	女	不转换	常量		-1.504	.260	-5.787	<.001
			AR, 季节性	延迟 1	.962	.221	4.361	<.001
				延迟 2	-.650	.153	-4.241	<.001
			季节性差异		1			
			MA, 季节性	延迟 1	1.475	.278	5.304	<.001
				延迟 2	-.611	.349	-1.754	.082
			时间	不转换	分子			
			延迟 0		.008	.003	2.305	.023

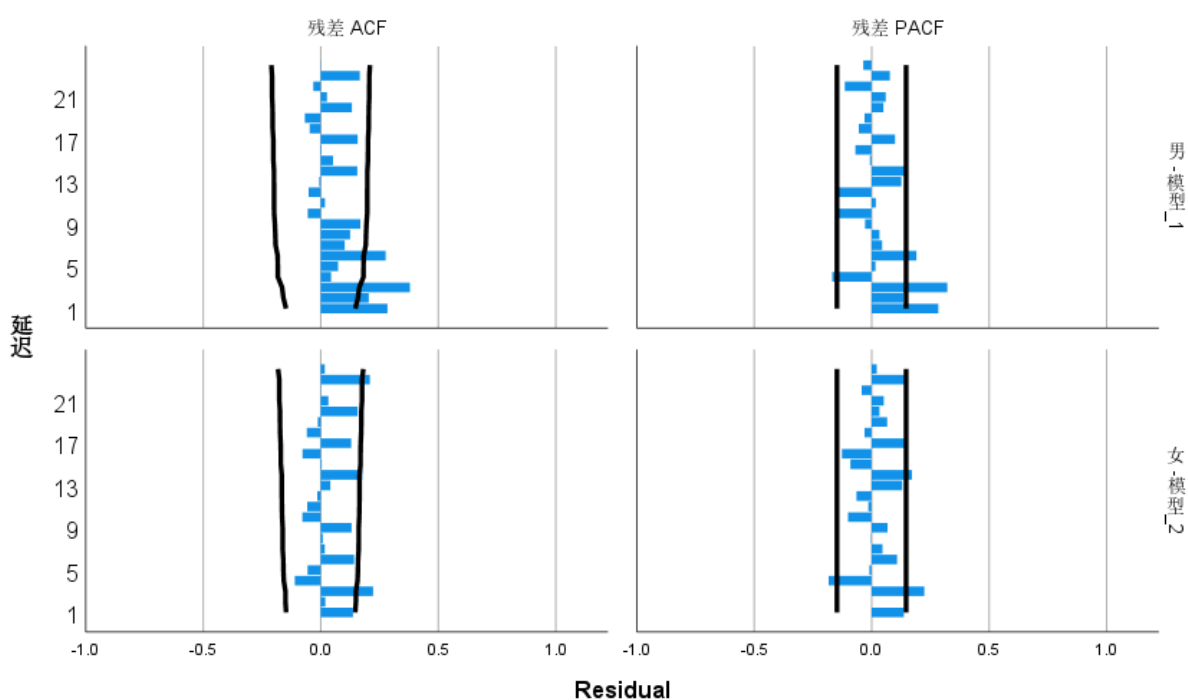
4.3.8 残差白噪声检验

尽管Q统计量很低，但在残差图可以验证两个模型的残差ACF均是白噪声。区别在于ARIMA模型的残差ACF极值很大，SARIMA模型的残差ACF分布较平缓。

图十三 ARIMA模型的残差ACF和残差PACF

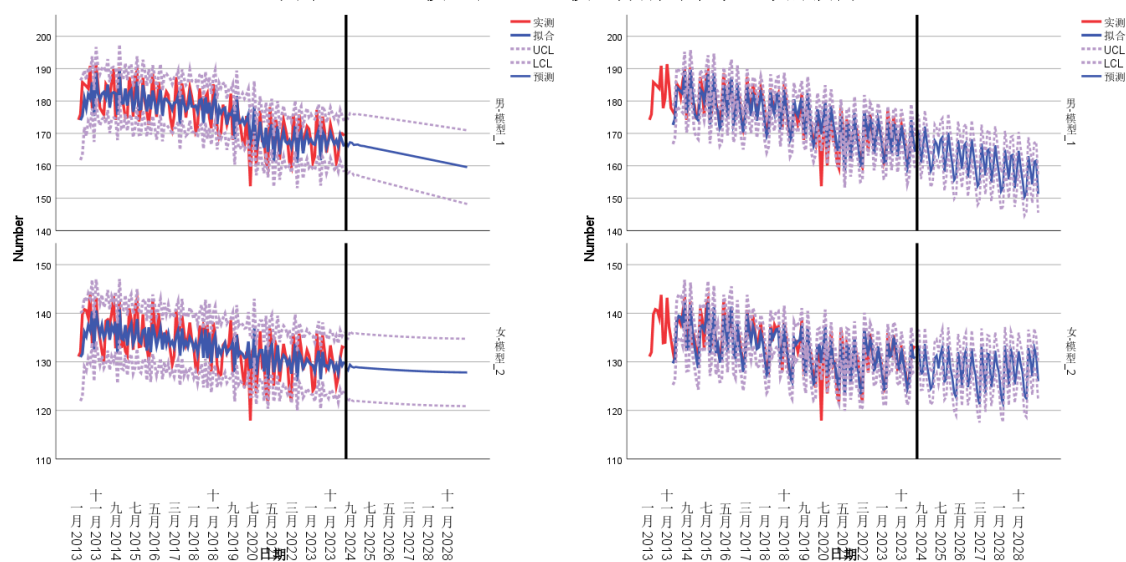


图十四 SARIMA模型的残差ACF和残差PACF



4.3.9 模型评估和预测

图十五 ARIMA模型和SARIMA模型分别对未来五年的预测



用两个模型分别对未来5年的月均工作时间进行预测，得到如图所示的图像。

从拟合效果来看，SARIMA模型对原始数据的拟合较好，因为ARIMA模型的数据走势过于平稳，不能很好地模拟出峰值，置信区间过宽，表示不确定性高。

经参数可靠性检验（略），其他参数所拟合的模型均没有（2，1，2）效果好。

五、结果的现实意义与建议

5.1 现实意义

首先，在对男女工作时长的性别差异和家庭角色差异分析中，我们统计发现：家庭角色不同，所带来的工作生活平衡是显著的。户主秩平均值最高。这是因为户主承担着全家收入来源的职责，平衡工作与生活的愿望也最强烈。而配偶、孩童和亲戚在家庭中则无需承担过多的职责，因此秩平均值最低。

其次，在时间序列分析中，我认为工作时间的季节性是显著的。从模型效果来看，两个模型在2023年之前的预测精度都较高，实际值和预测值的吻合度较好，而考虑季节性差异的SARIMA模型的拟合效果更好一些。从实际意义上看，这是由于日本每年春季固定举行的春斗所导致的。春斗，是劳工主要为了提高薪资与改善工作条件而发动的劳工运动，每年的春斗会依春季斗争的结果来提高薪资水准。从原始时间序列图中可以看出，极小值点都大部分出现在春季时期，与结论相符。

最后，ARIMA和SARIMA模型都预测了未来五年日本劳动者男女两性月均工作时间的减少。这表明劳动力市场可能正在经历结构性变化，可能由于人口老龄化、工作模式的变化（如远程办公）、以及劳动法的改革等。从预测图中也可以看出，尽管男性和女性的工作时间都在减少，且减少速率也比女性要快，但男性的月均工作时间通常比女性高，。这可能反映出传统的性别角色分工，男性仍然承担更多的工作时间。

政策制定者需要关注这一趋势，并探讨减少工作时间对经济生产力的影响，同时确保劳动者的生活质量和工作满意度。可以考虑如何在工作时间和家庭责任之间找到更好的平衡，以促进性别平等。例如，通过鼓励男性更多地参与家庭责任，或提供更多的灵活工作安排。

总的来说，这些预测显示出未来工作时间的减少趋势，这对于政策制定者和企业管理者都有重要的现实意义。通过合理的政策和管理措施，可以有效应对这一趋势，确保劳动者的福利和经济生产力的可持续发展。

5.2 合理化建议和改进建议

本报告忽略了“年龄”“职业”等因素对工作生活平衡的影响，后续研究可在此基础上改进。ARIMA模型的预测区间是基于残差不相关且服从正态分布的假设的，因此如果前提条件之一不被满足，预测区间就可能是错误的。本报告中的残差分布不是很好，可能会导致预测的误差偏大。

另外，对于大多数预测区间的计算而言，基于 ARIMA 的预测区间相对过于狭窄，这是因为只有误差的方差被加入考量。实际上，参数估计和模型阶数估计同样存在方差，而这些没有在计算时被加以考虑。除此之外，该方法假设建模时数据的历史特征会一直持续到未来，这也是预测区间过于狭窄的原因之一。

六、参考文献

- [1] Geneva: International Labour Office ,2022,Working Time and Work-Life Balance Around the World, ISBN: Print: 9789220379417. ISBN: Web PDF: 9789220378915
- [2] Clutterbuck, D. (2003). Managing the Work-Life Balance. *Journal of Management*, pp. 101-120.
- [3] 范围, 2014.1, 《日本劳动关系“三支柱”的形成、变革与展望》, 国家行政学院学报, DOI:10.14063/j.cnki.1008-9314.2014.01.016
- [4] Economist, August 2nd 2018, Death by work: Japan' s habits of overwork are hard to change, *Internet*.
- [5] 脳・心臓疾患の労災認定－「過労死」と労災保険－. 厚生労働省. [2014-02-14]
- [6] "Global, regional, and national burdens of ischemic heart disease and stroke attributable to exposure to long working hours for 194 countries, 2000–2016: A systematic analysis from the WHO/ILO Joint Estimates of the Work-related Burden of Disease and Injury". *Environment International*. 154: 106595. Bibcode:2021EnInt.15406595P. doi: 10.1016/j.envint.2021.106595. ISSN 0160-4120. PMC 8204267. PMID 34011457.
- [7] “ジェンダー平等 日本は146か国中で116位 男女間に大きな格差” . NHK (2022年7月13日). 2022年7月15日閲覧。
- [8] “【出口調査】衆院選の争点 10代は「新型コロナ対応」が最多（日本テレビ系（NNN））” . Yahoo!ニュース. 2021年11月2日閲覧。
- [9] Mark D. West*, December 12 2003 , Employment Market Institutions and Japanese Working Hours, *Law & Economics Working Papers Archive: 2003-2009, University of Michigan Law School*, Art. 22 [2003].
- [10] Hiroshi Ono, February 2018, Why do the Japanese work long hours? Sociological Perspectives on Long Working Hours in Japan. *Japan Labor Issues*, vol.2, no.5.
- [11] David Guest, June 2002, Perspectives on the Study of Work-Life Balance, *Social Science, Information* ,41(2):255-272.
- [12] Kwiatkowski, D., Phillips, P. C. B., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, 54(1-3), 159–178. [DOI]
- [13] Box, George; Jenkins, Gwilym (1970). Time Series Analysis: Forecasting and Control.

San Francisco: Holden-Day.

- [14]Kuniko Ishiguro ,Gender equality in the Japanese workplace : Development over the 35years since the enforcement of the EEOL, 文京学院大学外国語学部紀要 第21号 (2021)
- [15]Chie Aoyagi, WORK-LIFE BALANCE AND GENDER EQUALITY IN JAPAN, National Graduate Institute for Policy Studies.
- [16]Rie S. Hori, Progress and problems of gender equality in Japanese academics and geosciences, *Adv. Geosci.*, 53, 195–203, 2020.
- [17]Maiko Hata, Gender Inequality in the Japanese Workplace: Issues Related to the Promotion of Childcare Leave Taken by Men (2021), Gender Studies Program, Faculty of Humanities, Charles University, Prague.
- [18]闫祥祥, 使用 ARIMA模型预测公园绿地面积, 计算机科学 DOI:10.11896/jsjlx. 200300999
- [19] 王燕. 应用时间序列分析[M]. 北京: 中国人民大学出版社 2005.
- [20] 郝军章, 崔玉杰, 韩江雪. 基于SARIMA模型在我国铁路客运量中的预测[J]. 数学的实践与认识, 2015(18):10.
- [21]Brockwell, P. J., & Davis, R. A. (2016). *Introduction to time series and forecasting* (3rd ed). New York, USA: Springer.
- [22]Table 1: Annual Hours Worked, 1923-1939. Note: Rodosho, Rodo Tokei Chingin Maitsuki Chosa [Monthly Labor Survey]; Nichigin, Rodo Tokei [Labor Statistics], various years.
- [23]Table 2: Average number of days of paid leave per year used by workers in some countries/regions. Note: MISHA YURCHENKO, *Is Japanese work culture as bad as people say?* MAR 20, 2022 edition, Retrieved from: <https://mishayurchenko.substack.com/p/is-japanese-work-culture-as-bad-as>
- [24]Table 3: Relationship between working hours and productivity, OECD statistics, stats.oecd.org, 2015.

七、附录

7.1 ARIMA模型拟合度和残差数据

模型拟合度											
拟合统计	平均值	标准误差	最小值	最大值	百分位数						
					5	10	25	50	75	90	95
平稳 R 方	.552	.040	.523	.580	.523	.523	.523	.552	.580	.580	.580
R 方	.545	.133	.451	.639	.451	.451	.451	.545	.639	.639	.639
RMSE	4.273	.743	3.748	4.799	3.748	3.748	3.748	4.273	4.799	4.799	4.799
MAPE	2.158	.032	2.136	2.181	2.136	2.136	2.136	2.158	2.181	2.181	2.181
MaxAPE	9.110	.266	8.922	9.298	8.922	8.922	8.922	9.110	9.298	9.298	9.298
MAE	3.311	.589	2.895	3.728	2.895	2.895	2.895	3.311	3.728	3.728	3.728
MaxAE	12.338	1.945	10.962	13.713	10.962	10.962	10.962	12.338	13.713	13.713	13.713
正态化 BIC	3.106	.350	2.859	3.353	2.859	2.859	2.859	3.106	3.353	3.353	3.353

残差 ACF 摘要											
延迟	平均值	标准误差	最小值	最大值	百分位数						
					5	10	25	50	75	90	95
延迟 1	.030	.001	.029	.031	.029	.029	.029	.030	.031	.031	.031
延迟 2	.038	.031	.016	.060	.016	.016	.016	.038	.060	.060	.060
延迟 3	.038	.028	.018	.058	.018	.018	.018	.038	.058	.058	.058
延迟 4	.012	.022	-.003	.028	-.003	-.003	-.003	.012	.028	.028	.028
延迟 5	-.083	.012	-.091	-.074	-.091	-.091	-.091	-.083	-.074	-.074	-.074
延迟 6	.051	.058	.010	.092	.010	.010	.010	.051	.092	.092	.092
延迟 7	-.099	.010	-.106	-.091	-.106	-.106	-.106	-.099	-.091	-.091	-.091
延迟 8	-.013	.020	-.027	.001	-.027	-.027	-.027	-.013	.001	.001	.001
延迟 9	.064	.030	.043	.085	.043	.043	.043	.064	.085	.085	.085
延迟 10	-.230	.048	-.263	-.196	-.263	-.263	-.263	-.230	-.196	-.196	-.196
延迟 11	.085	.019	.071	.098	.071	.071	.071	.085	.098	.098	.098
延迟 12	.600	.009	.593	.606	.593	.593	.593	.600	.606	.606	.606
延迟 13	.024	.017	.012	.036	.012	.012	.012	.024	.036	.036	.036
延迟 14	-.034	.051	-.070	.002	-.070	-.070	-.070	-.034	.002	.002	.002
延迟 15	.046	.040	.017	.074	.017	.017	.017	.046	.074	.074	.074
延迟 16	-.061	.021	-.076	-.046	-.076	-.076	-.076	-.061	-.046	-.046	-.046
延迟 17	.008	.014	-.002	.018	-.002	-.002	-.002	.008	.018	.018	.018
延迟 18	-.053	.059	-.094	-.012	-.094	-.094	-.094	-.053	-.012	-.012	-.012
延迟 19	-.099	.007	-.104	-.094	-.104	-.104	-.104	-.099	-.094	-.094	-.094
延迟 20	-.008	.014	-.017	.002	-.017	-.017	-.017	-.008	.002	.002	.002
延迟 21	.047	.032	.024	.070	.024	.024	.024	.047	.070	.070	.070
延迟 22	-.287	.043	-.318	-.256	-.318	-.318	-.318	-.287	-.256	-.256	-.256
延迟 23	.147	.040	.119	.176	.119	.119	.119	.147	.176	.176	.176
延迟 24	.339	.018	.327	.352	.327	.327	.327	.339	.352	.352	.352

残差 PACF 摘要

延迟	平均值	标准误差	最小值	最大值	百分位数						
					5	10	25	50	75	90	95
延迟 1	.030	.001	.029	.031	.029	.029	.029	.030	.031	.031	.031
延迟 2	.038	.031	.016	.059	.016	.016	.016	.038	.059	.059	.059
延迟 3	.036	.027	.017	.055	.017	.017	.017	.036	.055	.055	.055
延迟 4	.008	.026	-.010	.026	-.010	-.010	-.010	.008	.026	.026	.026
延迟 5	-.087	.015	-.098	-.076	-.098	-.098	-.098	-.087	-.076	-.076	-.076
延迟 6	.055	.059	.013	.096	.013	.013	.013	.055	.096	.096	.096
延迟 7	-.097	.001	-.098	-.096	-.098	-.098	-.098	-.097	-.096	-.096	-.096
延迟 8	-.002	.014	-.012	.008	-.012	-.012	-.012	-.002	.008	.008	.008
延迟 9	.073	.036	.047	.098	.047	.047	.047	.073	.098	.098	.098
延迟 10	-.243	.057	-.283	-.202	-.283	-.283	-.283	-.243	-.202	-.202	-.202
延迟 11	.124	.013	.115	.133	.115	.115	.115	.124	.133	.133	.133
延迟 12	.641	.007	.636	.646	.636	.636	.636	.641	.646	.646	.646
延迟 13	-.034	.053	-.072	.003	-.072	-.072	-.072	-.034	.003	.003	.003
延迟 14	-.276	.029	-.297	-.255	-.297	-.297	-.297	-.276	-.255	-.255	-.255
延迟 15	.009	.006	.005	.013	.005	.005	.005	.009	.013	.013	.013
延迟 16	.011	.058	-.030	.052	-.030	-.030	-.030	.011	.052	.052	.052
延迟 17	.089	.051	.054	.125	.054	.054	.054	.089	.125	.125	.125
延迟 18	-.168	.014	-.178	-.158	-.178	-.178	-.178	-.168	-.158	-.158	-.158
延迟 19	-.041	.002	-.043	-.040	-.043	-.043	-.043	-.041	-.040	-.040	-.040
延迟 20	.024	.017	.012	.036	.012	.012	.012	.024	.036	.036	.036
延迟 21	.030	.001	.030	.031	.030	.030	.030	.030	.031	.031	.031
延迟 22	.041	.006	.037	.045	.037	.037	.037	.041	.045	.045	.045
延迟 23	.072	.038	.045	.099	.045	.045	.045	.072	.099	.099	.099
延迟 24	-.116	.033	-.139	-.092	-.139	-.139	-.139	-.116	-.092	-.092	-.092

7.2 SARIMA模型拟合度和残差数据

模型拟合度

拟合统计	平均值	标准误差	最小值	最大值	百分位数						
					5	10	25	50	75	90	95
平稳 R 方	.285	.058	.244	.325	.244	.244	.244	.285	.325	.325	.325
R 方	.815	.030	.794	.837	.794	.794	.794	.815	.837	.837	.837
RMSE	2.686	.646	2.229	3.143	2.229	2.229	2.229	2.686	3.143	3.143	3.143
MAPE	1.289	.100	1.218	1.360	1.218	1.218	1.218	1.289	1.360	1.360	1.360
MaxAPE	10.227	.258	10.045	10.409	10.045	10.045	10.045	10.227	10.409	10.409	10.409
MAE	1.978	.526	1.607	2.350	1.607	1.607	1.607	1.978	2.350	2.350	2.350
MaxAE	13.921	2.939	11.843	15.999	11.843	11.843	11.843	13.921	15.999	15.999	15.999
正态化 BIC	2.179	.486	1.835	2.522	1.835	1.835	1.835	2.179	2.522	2.522	2.522

残差 ACF 摘要

延迟	平均值	标准误差	最小值	最大值	百分位数						
					5	10	25	50	75	90	95
延迟 1	.211	.104	.138	.284	.138	.138	.138	.211	.284	.284	.284
延迟 2	.112	.131	.020	.205	.020	.020	.020	.112	.205	.205	.205
延迟 3	.301	.111	.223	.380	.223	.223	.223	.301	.380	.380	.380
延迟 4	-.033	.109	-.110	.044	-.110	-.110	-.110	-.033	.044	.044	.044
延迟 5	.009	.092	-.056	.074	-.056	-.056	-.056	.009	.074	.074	.074
延迟 6	.209	.095	.142	.276	.142	.142	.142	.209	.276	.276	.276
延迟 7	.060	.060	.017	.102	.017	.017	.017	.060	.102	.102	.102
延迟 8	.067	.082	.009	.125	.009	.009	.009	.067	.125	.125	.125
延迟 9	.150	.027	.131	.169	.131	.131	.131	.150	.169	.169	.169
延迟 10	-.066	.016	-.078	-.055	-.078	-.078	-.078	-.066	-.055	-.055	-.055
延迟 11	-.020	.054	-.058	.018	-.058	-.058	-.058	-.020	.018	.018	.018
延迟 12	-.034	.025	-.052	-.016	-.052	-.052	-.052	-.034	-.016	-.016	-.016
延迟 13	.017	.033	-.006	.041	-.006	-.006	-.006	.017	.041	.041	.041
延迟 14	.166	.014	.156	.176	.156	.156	.156	.166	.176	.176	.176
延迟 15	.029	.034	.005	.053	.005	.005	.005	.029	.053	.053	.053
延迟 16	-.037	.057	-.077	.004	-.077	-.077	-.077	-.037	.004	.004	.004
延迟 17	.144	.019	.130	.157	.130	.130	.130	.144	.157	.157	.157
延迟 18	-.053	.008	-.059	-.047	-.059	-.059	-.059	-.053	-.047	-.047	-.047
延迟 19	-.040	.039	-.067	-.012	-.067	-.067	-.067	-.040	-.012	-.012	-.012
延迟 20	.145	.018	.133	.157	.133	.133	.133	.145	.157	.157	.157
延迟 21	.030	.004	.027	.032	.027	.027	.027	.030	.032	.032	.032
延迟 22	-.015	.023	-.032	.001	-.032	-.032	-.032	-.015	.001	.001	.001
延迟 23	.188	.030	.167	.209	.167	.167	.167	.188	.209	.209	.209
延迟 24	.009	.012	.001	.018	.001	.001	.001	.009	.018	.018	.018

残差 PACF 摘要

延迟	平均值	标准误差	最小值	最大值	百分位数						
					5	10	25	50	75	90	95
延迟 1	.211	.104	.138	.284	.138	.138	.138	.211	.284	.284	.284
延迟 2	.068	.095	.001	.135	.001	.001	.001	.068	.135	.135	.135
延迟 3	.274	.069	.224	.323	.224	.224	.224	.274	.323	.323	.323
延迟 4	-.175	.011	-.183	-.168	-.183	-.183	-.183	-.175	-.168	-.168	-.168
延迟 5	.004	.018	-.009	.016	-.009	-.009	-.009	.004	.016	.016	.016
延迟 6	.150	.058	.109	.191	.109	.109	.109	.150	.191	.191	.191
延迟 7	.046	.001	.045	.046	.045	.045	.045	.046	.046	.046	.046
延迟 8	.015	.026	-.003	.034	-.003	-.003	-.003	.015	.034	.034	.034
延迟 9	.020	.068	-.028	.068	-.028	-.028	-.028	.020	.068	.068	.068
延迟 10	-.120	.028	-.140	-.100	-.140	-.140	-.140	-.120	-.100	-.100	-.100
延迟 11	.002	.022	-.014	.018	-.014	-.014	-.014	.002	.018	.018	.018
延迟 12	-.103	.055	-.142	-.065	-.142	-.142	-.142	-.103	-.065	-.065	-.065
延迟 13	.129	.003	.126	.131	.126	.126	.126	.129	.131	.131	.131
延迟 14	.160	.017	.148	.171	.148	.148	.148	.160	.171	.171	.171
延迟 15	-.048	.059	-.090	-.007	-.090	-.090	-.090	-.048	-.007	-.007	-.007
延迟 16	-.098	.041	-.127	-.069	-.127	-.127	-.127	-.098	-.069	-.069	-.069
延迟 17	.120	.028	.100	.140	.100	.100	.100	.120	.140	.140	.140
延迟 18	-.042	.018	-.055	-.030	-.055	-.055	-.055	-.042	-.030	-.030	-.030
延迟 19	.018	.069	-.031	.066	-.031	-.031	-.031	.018	.066	.066	.066
延迟 20	.042	.011	.034	.050	.034	.034	.034	.042	.050	.050	.050
延迟 21	.056	.006	.052	.060	.052	.052	.052	.056	.060	.060	.060
延迟 22	-.078	.050	-.114	-.043	-.114	-.114	-.114	-.078	-.043	-.043	-.043
延迟 23	.112	.049	.078	.146	.078	.078	.078	.112	.146	.146	.146
延迟 24	-.007	.041	-.036	.022	-.036	-.036	-.036	-.007	.022	.022	.022

7.3 寻找模型参数的差分检验代码示例

```
import pandas as pd
import matplotlib.pyplot as plt
from statsmodels.graphics.tsaplots import plot_acf, plot_pacf
from statsmodels.tsa.stattools
import adfuller

# 加载数据
file_path = r'D:\Personal\Desktop\thesis\data Analysis\时间序列
分析数据2.xlsx'
data = pd.read_excel(file_path)

# 确保时间列为日期时间格式
data['时间'] = pd.to_datetime(data['时间'], format='%b. %Y')

# 设置时间列为索引
data.set_index('时间', inplace=True)

# 提取 “平均每月工作小时” 列
monthly_hours_ts = data['平均每月工作小时']

# 定义一个函数来进行ADF检验并返回结果
def adf_test(series, name, max_diff=2, seasonal_diff=None):
    results = []
    original_series = series.copy()
```

```

for d in range(max_diff + 1):
    if d > 0:
        series = series.diff().dropna()

    adf_result = adfuller(series, autolag='AIC')
    t_stat = adf_result[0]
    p_value = adf_result[1]
    aic = adf_result[5]
    crit_values = adf_result[4]
    results.append({
        '变量': name,
        '差分阶数': d,
        't': t_stat,
        'P': p_value,
        'AIC': aic,
        '临界值 1%': crit_values['1%'],
        '临界值 5%': crit_values['5%'],
        '临界值 10%': crit_values['10%']
    })

# 季节性差分检验
if seasonal_diff:
    series_seasonal_diff =
original_series.diff(seasonal_diff).dropna()

```

```

adf_result = adfuller(series_seasonal_diff, autolag='AIC')
t_stat = adf_result[0]
p_value = adf_result[1]
aic = adf_result[5]
crit_values = adf_result[4]
results.append({
    '变量': name,
    '差分阶数': f'{seasonal_diff}季',
    't': t_stat,
    'P': p_value,
    'AIC': aic,
    '临界值 1%': crit_values['1%'],
    '临界值 5%': crit_values['5%'],
    '临界值 10%': crit_values['10%']
})

```

```

return results

```

进行ADF检验并生成结果表

```

adf_results = adf_test(monthly_hours_ts, '平均每月工作小时',
max_diff=2, seasonal_diff=12)

```

转换结果为DataFrame

```

adf_results_df = pd.DataFrame(adf_results)

```

```
# 显示结果表
print(adf_results_df)

# 设置图形尺寸
plt.figure(figsize=(12, 12))

# 绘制一阶差分后的ACF图
plt.subplot(211)
plot_acf(monthly_hours_ts.diff().dropna(), lags=40, ax=plt.gca())
plt.title('ACF for Average Monthly Work Hours (1st Difference)')

# 绘制一阶差分后的PACF图
plt.subplot(212)
plot_pacf(monthly_hours_ts.diff().dropna(), lags=40,
ax=plt.gca())
plt.title('PACF for Average Monthly Work Hours (1st Difference)')

# 显示图形
plt.tight_layout()
plt.show()
```

八、致谢

感谢沈思老师这一年以来的敦敦教诲和无私督促！

祝沈思老师改卷快乐，假期快乐，生活快乐:D

谢谢！