

浙江工商大學

课 程 设 计



题目： 基于 VAR 模型的中国人口动态预测研究

课程名称：应用时间序列分析

学 院：统计与数学学院

专 业：应用统计

班 级：应统二班

学 号：22020040149

学生姓名：王振宽

二〇二三年 五 月

摘要

在本研究中，我们采用 VAR 模型对中国人口数据进行预测。数据集包含了六个变量，分别是人口密度（Population Density）、预期寿命（Life Expectancy）、出生率（Birth Rate）、死亡率（Death Rate）、婴儿死亡率（Infant Mortality Rate）和生育率（Fertility Rate）。

首先，我们对每个变量进行可视化分析，并发现随时间的变化趋势近似为线性关系且变量间存在高度相关性。接着，我们使用格兰杰因果关系检验来探究变量之间的因果关系。结果显示系统中的所有变量（时间序列）都存在相互影响的关系。为了进一步分析变量之间的长期稳定关系，我们进行了协整检验。然而，我们发现婴儿死亡率和生育率之间可能不存在明显的长期稳定关系。因为 ADF 检验未通过，为了获取稳定的序列，我们对数据进行了四阶差分。

为了确定 VAR 模型的适当阶数，我们逐步拟合了不同阶数的 VAR 模型，并选择具有最小 AIC 的模型阶数为滞后 6 阶。我们还使用德宾沃森统计量检验了残差（误差）序列的序列相关性，并发现其值接近 2，表明不存在自相关性。

最后，我们对预测结果进行了四次反转，得到最终的预测结果。通过评估指标，我们发现预测值与实际值非常接近，表明我们的模型效果良好。

综上所述，本研究采用 VAR 模型对中国人口数据进行了准确的预测。我们通过分析数据的变化趋势、检验因果关系、协整检验和差分操作，建立了有效的 VAR 模型，并通过评估指标验证了预测结果的准确性。这些研究结果对于理解中国人口发展和制定相关政策具有重要意义。

关键词：VAR，人口数据，格兰杰因果检验，ADF 检验

一、 引言

研究中国的人口数据是一个广泛而重要的领域，考察中国人口数据可以提供关于中国人口规模、结构、分布和变化的深入洞察。

近年来，中国人口数据的背景呈现出一些重要的特点和趋势：

人口总量稳定增长：尽管中国实施了计划生育政策，但由于人口基数大，近年来中国的人口总量仍然保持稳定增长。根据最新统计数据，中国的人口总量超过 14 亿人，仍是世界上人口最多的国家。

人口结构逐渐老龄化：随着长期实施计划生育政策以及社会经济发展和医疗水平提高，中国的人口结构呈现出老龄化的趋势。老年人口比例增加，劳动年龄人口比例相对减少，这对社会养老、医疗等方面提出了挑战。

城市化进程加速：中国的城市化进程快速推进，大量农村人口涌入城市。这导致城市人口比例的增加和城市人口的集中增长，对城市规划、基础设施建设、公共服务等提出了压力和需求。

生育率下降：随着计划生育政策的实施和社会发展的进步，中国的生育率持续下降。近年来，生育率继续保持在相对较低的水平，这对人口结构、劳动力供给和社会福利等方面产生了影响。

农村人口减少：农村地区的人口数量逐渐减少，这主要是由于农村人口向城市迁移以及农业现代化进程的推动。这对农村经济发展、农村社会服务和农村地区的可持续发展提出了新的挑战。

这些背景特点和趋势对于了解中国的人口状况、社会发展和政策制定具有重要意义，为相关领域的研究和规划提供了基础数据和参考依据。同时，它们也引发了一系列社会经济问题和挑战，需要适应性的政策和措施来应对。研究人口数据是一项很有意义的工作，以下是研究中国人口数据的一些背景和意义：

- 人口规模和趋势：**中国是世界上人口最多的国家之一，了解中国的人口规模和趋势对于了解全球人口格局至关重要。研究中国的人口数据可以揭示中国人口数量的变化趋势、年龄结构的演变以及城乡人口比例的变化，为社会政策和规划提供依据。
- 人口结构和性别比例：**研究中国的人口数据可以帮助了解人口的性别比例、年龄结构和教育水平等方面的变化。这对于评估社会需求、劳动力市

场和教育资源的分配具有重要意义。

- 3. 人口流动和城市化：中国的人口流动和城市化是一个重要的研究领域。通过分析人口迁移、农村转城市人口的变化以及城市化进程的影响，可以深入了解中国城市化的挑战和机遇，为城市规划和城市发展提供参考。
- 4. 健康和社会福利：人口数据研究对于评估中国的健康状况和社会福利具有重要意义。通过观察出生率、死亡率、预期寿命以及疾病流行情况等指标，可以评估中国的医疗保健系统和公共卫生政策的有效性，为改善人民健康和社会福利提供依据。
- 5. 人口政策和社会发展：中国的人口政策和计划生育政策对于人口数据研究具有重要影响。通过研究人口数据，可以评估人口政策的实施效果、人口控制的成就和挑战，为未来的人口政策和社会发展提供经验和参考。

总之，研究中国的人口数据可以帮助我们了解中国人口的规模、结构和变化趋势，揭示社会经济发展和政策制定的关键问题，为制定相关政策和规划提供科学依据。同时，这些研究也有助于了解中国的社会福利、健康状况和城市化进程，为促进可持续发展和改善人民生活水平做出贡献。

二、 数据集介绍

我们的数据集收集从 1960 年到 2021 年，指标有人口密度、预期寿命、出生率、死亡率、婴儿死亡率和生育率等六个指标。我们的部分数据集如下所示：

Year	Population Density	Life Expectancy	Birth Rate	Death Rate	Infant Mortality Rate	Fertility Rate
1960	68.14	44.5	37.629	21.031	132.647	5.748
1961	68.26	44.52	38.223	20.905	133.444	5.882
1962	69.23	44.53	38.818	20.779	134.241	6.016
1963	71.24	44.55	39.412	20.653	135.038	6.15
1964	73.4	46.73	39.427	19.082	126.795	6.18
1965	75.4	48.92	39.442	17.511	118.552	6.21
1966	77.39	51.1	39.458	15.941	110.31	6.24
1967	79.27	53.29	39.473	14.37	102.067	6.27
1968	81.29	55.47	39.488	12.799	93.824	6.3
1969	83.48	56.71	37.976	12.075	89.449	6.01
1970	85.68	57.95	36.463	11.351	85.073	5.72
1971	87.84	59.19	34.951	10.626	80.698	5.43
1972	89.88	60.43	33.438	9.902	76.322	5.14
1973	91.84	61.67	31.926	9.178	71.947	4.85
1974	93.68	62.44	30.039	8.772	68.557	4.482
1975	95.33	63.21	28.152	8.366	65.167	4.114
1976	96.81	63.97	26.265	7.959	61.778	3.746

图 1：数据集部分展示

这些人口数据指标在人口研究中具有重要的实际意义，下面我将逐个解释它们的意义：

- 1. 人口密度（Population Density）：人口密度是指单位面积或单位空间内的人口数量。它可以帮助我们了解人口在特定地区的分布情况，进而对城市规划、土地利用、基础设施建设等方面进行决策和规划。高人口密度可能

导致资源压力和环境负担，而低人口密度可能面临人口老龄化和发展不均衡的问题。

2. **预期寿命 (Life Expectancy):** 预期寿命是指在出生时预计一个人在特定地区平均能够活到的年龄。它是衡量人口健康水平和医疗保健状况的重要指标。较高的预期寿命通常与较好的医疗服务、营养状况和生活条件有关，可以帮助评估社会发展水平和改善人口福利的效果。
3. **出生率 (Birth Rate):** 出生率是指在特定地区每年每千人口中新生儿的数量。它反映了人口增长的趋势和水平，对于人口规划和社会经济发展具有重要影响。较高的出生率可能导致人口过快增长和人口结构失衡，而较低的出生率可能面临人口老龄化和劳动力供应不足的问题。
4. **死亡率 (Death Rate):** 死亡率是指在特定地区每年每千人口中死亡的人口数量。它可以用来评估一个地区的健康状况和医疗水平。较高的死亡率可能与疾病、营养不良、战争等因素有关，而较低的死亡率通常与良好的卫生保健和医疗条件有关。
5. **婴儿死亡率 (Infant Mortality Rate):** 婴儿死亡率是指在婴儿出生后的第一年内每千活产婴儿中死亡的人口数量。它是衡量一个地区婴儿健康状况和医疗保健质量的重要指标。较高的婴儿死亡率通常与贫困、疾病、不良卫生条件等因素有关，而较低的婴儿死亡率意味着较好的健康保健和生活条件。
6. **生育率 (Fertility Rate):** 生育率是指在特定年龄范围内平均每位女性的生育子女数量。它对于人口增长和人口结构的变化具有重要影响。较高的生育率可能导致人口增长速度加快和年龄结构年轻化，而较低的生育率可能导致人口老龄化和人口减少。

这些人口数据指标对于制定人口政策、规划社会经济发展、改善医疗保健和教育条件等方面都具有指导意义。通过监测和分析这些指标的变化，可以帮助决策者了解人口的趋势和变化，从而制定相应的措施和政策，促进社会的可持续发展。

三、 模型的建立

3.1. 模型的选择

时间序列问题是指对一系列按时间顺序排列的数据进行分析、建模和预测的问题。这些数据通常是按照固定时间间隔采集或记录的，可以是经济数据、金融数据、气象数据、人口数据等。

时间序列分析的目的是揭示数据中存在的模式、趋势、周期性和随机性，并利用这些信息进行预测和决策。时间序列分析方法可以帮助我们理解数据的动态特征、发现潜在规律、预测未来走势，并进行相关的政策制定、经济预测、市场调整等。

VAR（Vector Autoregression）模型是一种常用的经济计量模型，用于分析多个经济变量之间的相互关系和动态变化。VAR 模型是基于时间序列数据的统计模型，它假设变量之间存在互相关联的关系，并且可以相互影响和解释彼此的变化。

VAR 模型的基本思想是将多个经济变量组合成一个向量，通过引入滞后项来捕捉变量之间的动态关系。模型中的每个变量都被认为是其他变量的函数，包括它自己的滞后项和其他变量的滞后项。这样，VAR 模型能够描述变量之间的联动效应和相互影响，揭示它们的动态调整过程。

VAR 模型的建立需要进行一系列的步骤，包括模型阶数的选择、参数估计和模型诊断等。通常使用最小二乘法或极大似然估计等方法对 VAR 模型进行参数估计。模型的诊断包括残差的自相关性检验、稳定性检验和模型拟合程度的评估等，以验证模型的合理性和可靠性。

VAR 模型在经济学和金融学等领域广泛应用，可以用于预测经济变量的未来走势、评估政策效果、分析经济冲击的传导机制等。它能够提供对多个变量之间关系的全面理解，帮助研究人员更好地理解 and 解释经济现象的复杂性。同时，VAR 模型也有一些局限性，如数据的选择和模型的设定可能对结果产生影响，需要结合实际情况进行合理解释和使用。

在时间序列分析中选择使用 VAR 模型可以有以下考虑：

1. 多个变量之间的相互关系：VAR 模型适用于分析多个相关变量之间的相互影响和动态调整过程。如果你的研究对象包含多个经济变量或其他相

关变量，并且你希望了解它们之间的关系和相互影响，那么 VAR 模型是一个合适的选择。

2. 变量之间的内生性：VAR 模型假设变量之间存在内生性，即它们相互影响并被彼此解释。如果你认为你的变量是内生的，即它们的变化不仅受到外部因素的影响，还受到系统内部其他变量的影响，那么 VAR 模型可以帮助你捕捉这种内在关系。
3. 动态特征和时序依赖性：VAR 模型能够捕捉变量之间的动态特征和时序依赖性，它考虑了变量的滞后项，并可以揭示变量之间的动态调整过程。如果你关注变量的短期和长期动态特征，并且认为它们的变化是相互依赖和连续的，那么 VAR 模型可以提供有关这些动态特征的信息。
4. 预测和冲击分析：VAR 模型可以用于预测未来的变量值，并评估不同变量之间的冲击传导机制。如果你希望对未来的变量走势进行预测，并研究不同变量之间的冲击效应和传导机制，那么 VAR 模型是一个有用的工具。

然而，选择 VAR 模型也需要注意一些限制和假设，例如模型的稳定性、滞后阶数的选择、外生变量的考虑等。在使用 VAR 模型之前，建议进行模型诊断和敏感性分析，以确保模型的合理性和可靠性。此外，对于某些特定问题，其他时间序列模型（如 ARIMA、GARCH 等）或者结构方程模型（如 VARMA、VARX 等）也可能是更合适的选择。因此，在选择模型时，需要结合具体问题和数据特点进行综合考虑。

对于研究人口数据的指标，如 Population Density（人口密度）、Life Expectancy（预期寿命）、Birth Rate（出生率）、Death Rate（死亡率）、Infant Mortality Rate（婴儿死亡率）和 Fertility Rate（生育率），VAR 模型可以是一个不错的选择。VAR 模型（Vector Autoregression）是一种用于分析多个变量之间动态关系的时间序列模型。它通常用于描述和预测变量之间的相互依赖关系，适用于具有较短时间跨度、高频率观测的数据。VAR 模型假设各个变量是内生的，即彼此相互影响，而对于人口数据中的指标来说，它们往往是由不同的原因和机制所驱动，并存在一些因果关系。

3.2. 向量自回归模型 (Vector Autoregression)

VAR 中的所有变量都以相同的方式进入模型：每个变量都有一个方程式，根据其自身的滞后值、其他模型变量的滞后值和一个误差项来解释其演变。向量自回归 (VAR) 是一种多变量预测算法，当两个或多个时间序列相互影响时使用。这意味着，使用 VAR 的基本要求是：

1. 至少需要两个时间序列 (变量)；
2. 时间序列应相互影响。

它被认为是自回归模型，因为每个变量 (时间序列) 都建模为过去值的函数，即预测变量只不过是序列的滞后 (时间延迟值)。那么 VAR 与其他自回归模型 (如 AR, ARMA 或 ARIMA) 有何不同？

主要区别在于这些模型是单向的，其中预测变量影响 Y ，反之亦然。而向量自动回归 (VAR) 是双向的。也就是说，变量相互影响。

在自回归模型中，时间序列被建模为其自身滞后的线性组合。也就是说，序列的过去值用于预测当前和未来。

典型的 AR (p) 模型方程如下所示：

$$Y_t = \alpha + \beta_1 Y_{t-1} + \beta_2 Y_{t-2} + \cdots + \beta_p Y_{t-p} + \varepsilon_t$$

其中 α 是截距，一个常数， β_1, β_2 直到 β_p 是 Y 到 p 阶的滞后的系数。阶 p 表示，最多使用 Y 的 p 滞后，它们是方程中的预测因子。 ε_t 是错误，被视为白噪声。

在 VAR 模型中，每个变量都被建模为自身过去值和系统中其他变量过去值的线性组合。由于我们有多组相互影响的时间序列，因此将其建模为方程组，每个变量 (时间序列) 一个方程。也就是说，如果有 5 个相互影响的时间序列，我们将拥有一个由 5 个方程组成的系统。

那么，方程是如何精确构建的？假设我们有两个变量 (时间序列) Y_1 和 Y_2 ，需要在时间 t 预测这些变量的值。为了计算 $Y_1(t)$ ，VAR 将使用 Y_1 和 Y_2 的过去值。同样，要计算 $Y_2(t)$ ，请使用 Y_1 和 Y_2 的过去值。

例如，具有两个时间序列 (变量 “ Y_1 ” 和 “ Y_2 ”) 的 VAR (2) 模型的方程组如下：

$$\begin{aligned} Y_{1,t} &= \alpha_1 + \beta_{11,1}Y_{1,t-1} + \beta_{12,1}Y_{2,t-1} + \varepsilon_{1,t} \\ Y_{2,t} &= \alpha_2 + \beta_{21,1}Y_{1,t-1} + \beta_{22,1}Y_{2,t-1} + \varepsilon_{2,t} \end{aligned}$$

其中， $Y_{1,t-1}$ 和 $Y_{2,t-1}$ 分别是时间序列 Y_1 和 Y_2 的第一个滞后。

上述方程称为 VAR (1) 模型，因为每个方程的阶数为 1，也就是说，它最多包含一个预测变量（ Y_1 和 Y_2 ）的滞后。由于方程中的 Y 项是相互关联的，因此 Y 被视为内生变量，而不是外生预测因子。类似地，两个变量的二阶 VAR (2) 模型将为每个变量（ Y_1 和 Y_2 ）包含最多两个滞后。

$$\begin{aligned} Y_{1,t} &= \alpha_1 + \beta_{11,1}Y_{1,t-1} + \beta_{12,1}Y_{2,t-1} + \beta_{11,2}Y_{1,t-2} + \beta_{12,2}Y_{2,t-2} + \varepsilon_{1,t} \\ Y_{2,t} &= \alpha_2 + \beta_{21,1}Y_{1,t-1} + \beta_{22,1}Y_{2,t-1} + \beta_{21,2}Y_{1,t-2} + \beta_{22,2}Y_{2,t-2} + \varepsilon_{2,t} \end{aligned}$$

能写出一个有三个变量（ Y_2 、 Y_1 和 Y_3 ）的二阶 VAR(3)模型会是：

$$\begin{aligned} Y_{1,t} &= \alpha_1 + \beta_{11,1}Y_{1,t-1} + \beta_{12,1}Y_{2,t-1} + \beta_{13,1}Y_{3,t-1} + \beta_{11,2}Y_{1,t-2} + \beta_{12,2}Y_{2,t-2} + \beta_{13,2}Y_{3,t-2} + \varepsilon_{1,t} \\ Y_{2,t} &= \alpha_2 + \beta_{21,1}Y_{1,t-1} + \beta_{22,1}Y_{2,t-1} + \beta_{23,1}Y_{3,t-1} + \beta_{21,2}Y_{1,t-2} + \beta_{22,2}Y_{2,t-2} + \beta_{23,2}Y_{3,t-2} + \varepsilon_{2,t} \\ Y_{3,t} &= \alpha_3 + \beta_{31,1}Y_{1,t-1} + \beta_{32,1}Y_{2,t-1} + \beta_{33,1}Y_{3,t-1} + \beta_{31,2}Y_{1,t-2} + \beta_{32,2}Y_{2,t-2} + \beta_{33,2}Y_{3,t-2} + \varepsilon_{3,t} \end{aligned}$$

随着模型中时间序列（变量）数量的增加，方程组会变大。

构建 VAR 模型的过程涉及以下步骤：

1. 分析时间序列特征；
2. 检验时间序列之间的因果关系；
3. 平稳性检验；
4. 如果需要，转换序列以使其稳定；
5. 找到最佳顺序 p ；
6. 准备训练和测试数据集；
7. 训练模型；
8. 转换回差分前的数据（如果有）；
9. 使用测试集评估模型；
10. 对未来的预测。

四、模型的求解

4.1 分析时间序列特征

我们的每一列都是时间序列，首先看一下变量之间的相关性分析，如下图所示：

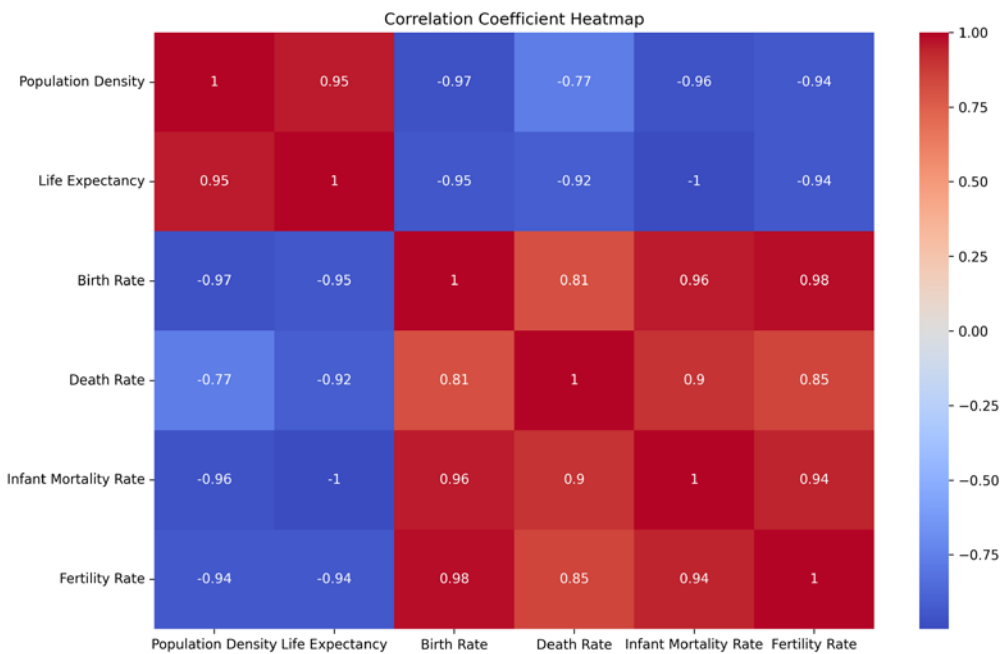
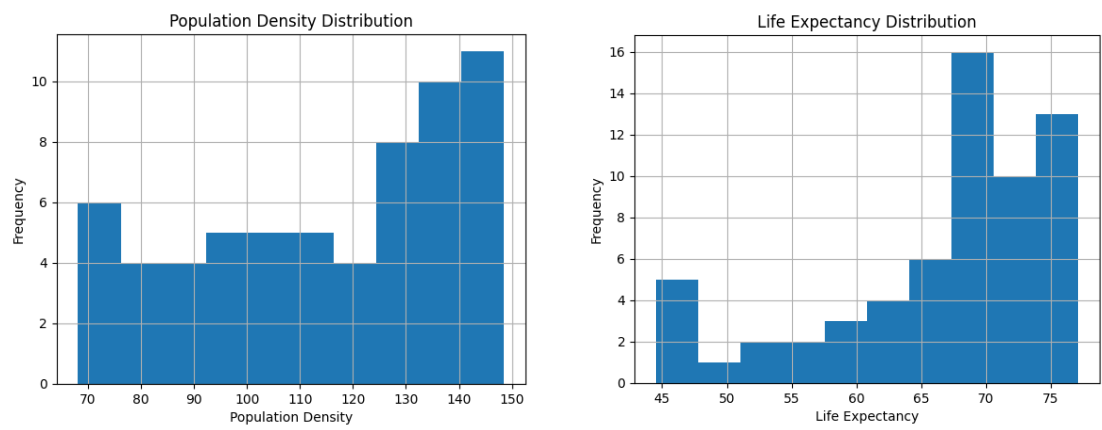


图 2：变量之间的相关性热力图

生成一个相关系数热力图，显示了数据框中各列之间的相关性程度。颜色的深浅表示相关系数的大小，正值表示正相关，负值表示负相关。从上图中不难看出，变量之间都有着高度相关性。

我们看一下每个变量的直方图情况，如下所示：



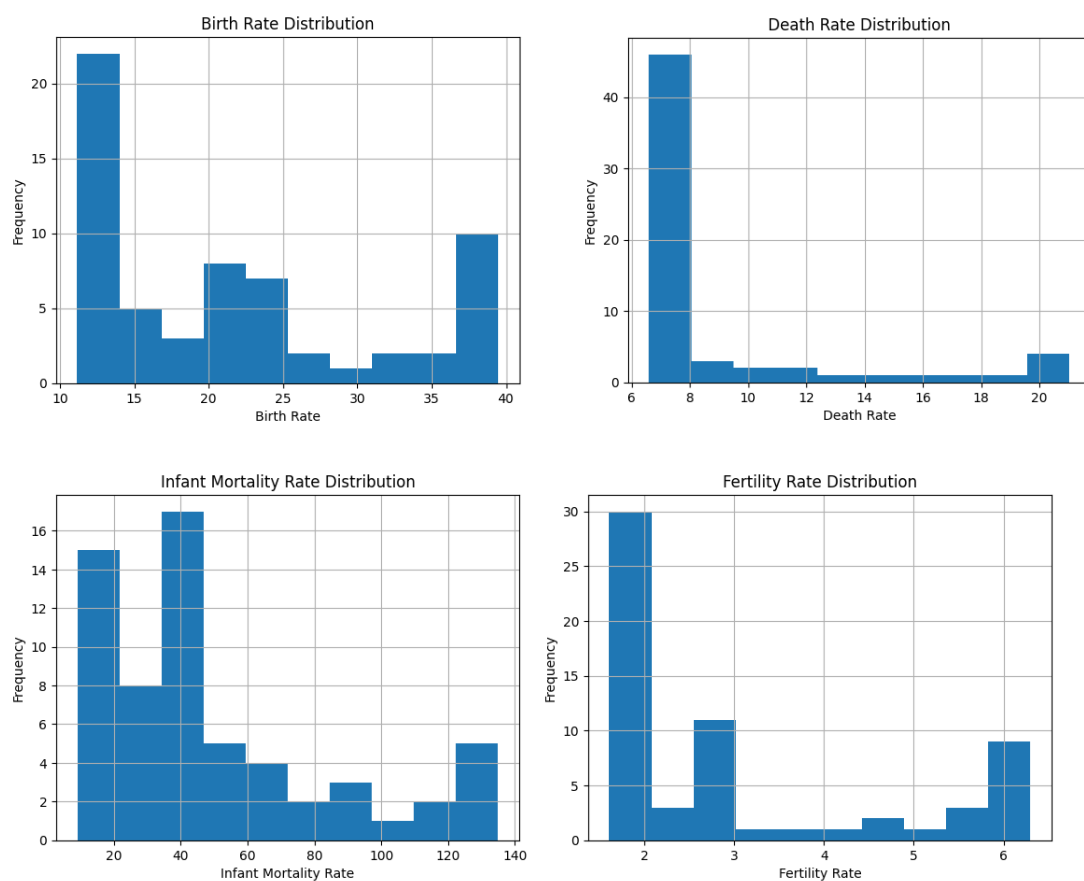


图 3: 六个变量的直方图

可以看到，我们的数据不符合线性变化，且分布左偏。

接下来画出六个时间序列的图像，如下图所示：



图 4: 六个时间序列的图像

时间序列图可以将数据随时间的变化趋势和模式直观地展示出来。通过观察图像，可以识别出数据中的长期趋势、季节性模式、周期性变化以及其他重要的模式。这有助于揭示数据背后的规律和趋势。可以看到，除了人口密度和期望的寿命长度两个变量随着时间的推移而增大，其他变量都是随着时间的推移而减小。

虽然部分变量的变化不是线性的，但是变化趋势还是可以通过线性来表示。故我们选择 VAR 模型是有效可行的。

4.2 格兰杰因果检验

格兰杰因果检验（Granger causality test）是一种用于分析时间序列数据中的因果关系的统计检验方法。它基于时间序列数据的自回归模型，通过比较包含和不包含滞后项的模型来判断一个变量是否能够因果地影响另一个变量。

在格兰杰因果检验中，我们考虑两个变量： X 和 Y 。该检验的原假设（null hypothesis）是 X 对 Y 没有因果影响，即 X 的过去值不能提供关于 Y 未来值的额外信息。如果原假设被拒绝，即 p 值小于预设显著性水平（通常为 0.05），则可以得出结论认为 X 对 Y 存在因果影响。

我们得到的格兰杰因果检验的结果如下表所示：

	Population Density_x	Life Expectancy_x	Birth Rate_x	Death Rate_x	Infant Mortality Rate_x	Fertility Rate_x
Population Density_y	1.0	0.0001	0.0	0.0	0.0001	0.0
Life Expectancy_y	0.0	1.0000	0.0	0.0	0.0000	0.0
Birth Rate_y	0.0	0.0000	1.0	0.0	0.0001	0.0
Death Rate_y	0.0	0.0000	0.0	1.0	0.0000	0.0
Infant Mortality Rate_y	0.0	0.0000	0.0	0.0	1.0000	0.0
Fertility Rate_y	0.0	0.0000	0.0	0.0	0.0000	1.0

图 5：格兰杰因果检验的结果

在输出结果中，index 以 y 结尾，表示响应变量，column 以 x 结尾，表示预测变量，如果 p 值小于 0.05 表明存在格兰杰因果性。因此根据检验数据，完全有理由使用 VAR 模型。

4.3 协整性检验

协整性检验在时间序列分析中具有重要的意义。它用于判断多个时间序列变量之间是否存在长期稳定的关系，即它们在长期内共同演化或共同受到某种平衡力量的影响。协整性检验的意义如下：

1. 发现长期关系：协整性检验可以帮助我们发现多个时间序列变量之间的

长期关系。相比于短期关系，长期关系更能反映经济、金融和其他领域中的结构性联系。通过协整性检验，我们可以确定变量之间是否存在长期均衡关系。

- 2. 提供有效的投资策略：协整性检验对于金融领域中的配对交易和套利策略具有重要意义。如果两个变量存在协整关系，即它们在长期内保持均衡状态，那么当它们的价格发生偏离时，我们可以考虑采取相应的交易策略以获取利润。
- 3. 避免伪回归：在时间序列数据中，变量之间可能存在表面上的相关性，但实际上是由于它们都受到相同的趋势或季节性影响。协整性检验可以帮助我们区分真正的长期关系和伪回归（spurious regression），从而避免错误的推断和分析。
- 4. 提供经济解释：协整性检验的结果可以提供对经济关系的解释。如果变量之间存在协整关系，我们可以通过协整向量（cointegration vector）的系数来理解它们之间的经济意义和影响机制。

综上所述，协整性检验在时间序列分析中具有重要的意义，它可以帮助我们发现长期关系、制定有效的投资策略、避免伪回归，并提供经济解释。它是进行多变量分析和时间序列建模的重要工具之一。

协整性检验的结果如下所示：

Name	::	Test Stat	> C(95%)	=>	Signif

Population Density	::	277.06	> 83.9383	=>	True
Life Expectancy	::	178.71	> 60.0627	=>	True
Birth Rate	::	91.03	> 40.1749	=>	True
Death Rate	::	30.76	> 24.2761	=>	True
Infant Mortality Rate	::	10.13	> 12.3212	=>	False
Fertility Rate	::	0.26	> 4.1296	=>	False

图 6：协整性检验的结果

每个变量名称后面的箭头（=>）表示该变量的检验统计量是否大于临界值。如果检验统计量大于临界值，则可以拒绝原假设，即存在协整关系。如果检验统计量小于等于临界值，则无法拒绝原假设，即不存在协整关系。

根据得到的结果，我们可以进行以下分析：

1. "Population Density"、"Life Expectancy"、"Birth Rate"和"Death Rate"的检验统计量都大于临界值，这表明在 95%的置信水平下，这些变量之间存在协整关系。这意味着它们之间存在长期稳定的关系，可能彼此相互影响。
2. "Infant Mortality Rate"和"Fertility Rate"的检验统计量都小于临界值，无法拒绝原假设，即在 95%的置信水平下，这些变量之间不存在协整关系。这意味着它们之间可能没有明显的长期稳定关系。

基于以上分析，我们可以得出结论：

1. "Population Density"、"Life Expectancy"、"Birth Rate"和"Death Rate"之间可能存在长期稳定的关系，可以考虑进一步研究它们之间的因果关系和影响机制。
2. "Infant Mortality Rate"和"Fertility Rate"之间可能没有明显的长期稳定关系，可能需要更多的数据或其他变量来进行进一步分析。

4.4 ADF 检验

我们先将原始数据集划分为训练集和测试集。训练集是原始数据集中的前 58 个观测值，用于训练模型。它的形状为(58, 6)，表示包含 58 行和 6 列的数据框（假设有 6 个变量）。测试集是原始数据集中的最后 4 个观测值，用于测试模型的性能。它的形状为(4, 6)，表示包含 4 行和 6 列的数据框。

通过将数据集划分为训练集和测试集，我们可以使用训练集来拟合模型，并使用测试集评估模型在新数据上的表现。这有助于验证模型的预测能力和泛化能力，以确保模型对未来数据的预测准确性。

ADF 测试（Augmented Dickey-Fuller test）是一种用于检验时间序列数据是否具有单位根（unit root）的统计检验方法。单位根表示时间序列数据存在非平稳性，即数据的均值和方差不随时间保持恒定。ADF 测试的目的是确定时间序列数据是否具有趋势性，并提供对数据平稳性的判断。

ADF 测试的原假设（null hypothesis）是时间序列数据具有单位根，即数据是非平稳的。如果原假设被拒绝，即 p 值小于预设显著性水平（通常为 0.05），则可以得出结论认为时间序列数据是平稳的。

ADF 测试基于自回归模型（Autoregressive model），其中包含了时间序列数

据的滞后项。检验的过程涉及以下步骤：

1. 将时间序列数据拆分为训练集和测试集。
2. 使用训练集数据拟合自回归模型。
3. 根据模型的拟合结果进行模型选择和参数估计。
4. 使用测试集数据计算 ADF 统计量。
5. 基于 ADF 统计量计算 p 值。
6. 判断 p 值是否小于显著性水平，从而判断是否拒绝原假设。

ADF 测试常用于时间序列分析中，特别是用于判断一个时间序列数据是否具有趋势性。它可以帮助我们确定数据的平稳性，并在建立模型、进行预测和进行统计推断时提供有关时间序列数据的重要信息。

我们得到的结果为：

Augmented Dickey-Fuller Test on "Population Density"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.5638

No. Lags Chosen = 8

Critical value 1% = -3.571

Critical value 5% = -2.923

Critical value 10% = -2.599

=> P-Value = 0.1007. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Life Expectancy"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -0.6138

No. Lags Chosen = 6

Critical value 1% = -3.566

Critical value 5% = -2.92

Critical value 10% = -2.598

=> P-Value = 0.8678. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Birth Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.6418

No. Lags Chosen = 11

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.0847. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Death Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.2498

No. Lags Chosen = 11

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.1887. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Infant Mortality Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -1.0009

No. Lags Chosen = 6

Critical value 1% = -3.566

Critical value 5% = -2.92

Critical value 10% = -2.598

=> P-Value = 0.7529. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Fertility Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -3.4284

No. Lags Chosen = 11

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.01. Rejecting Null Hypothesis.

=> Series is Stationary.

根据 ADF 检验的结果，对每个变量进行分析如下：

"Population Density": P-Value 为 0.1007，大于显著性水平 0.05，因此无足够证据拒绝原假设，该变量是非平稳的。

"Life Expectancy": P-Value 为 0.8678，大于显著性水平 0.05，因此无足够证据拒绝原假设，该变量是非平稳的。

"Birth Rate": P-Value 为 0.0847，略大于显著性水平 0.05，但仍然较接近，存在一定程度上的非平稳性的迹象。

"Death Rate": P-Value 为 0.1887，大于显著性水平 0.05，因此无足够证据拒绝原假设，该变量是非平稳的。

"Infant Mortality Rate": P-Value 为 0.7529，远远大于显著性水平 0.05，因此无足够证据拒绝原假设，该变量是非平稳的。

"Fertility Rate": P-Value 为 0.01, 小于显著性水平 0.05, 因此有足够证据拒绝原假设, 该变量是平稳的。

综上所述, 除了"Fertility Rate"变量是平稳的外, 其他变量均表现出非平稳的特征。在时间序列分析中, 平稳性是一个重要的前提条件, 因此在进一步的分析和建模中, 需要对非平稳变量进行处理, 例如进行差分操作或者使用合适的时间序列模型来处理非平稳性。

我们第一次差分操作后的结果为:

Augmented Dickey-Fuller Test on "Population Density"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -1.9854

No. Lags Chosen = 7

Critical value 1% = -3.571

Critical value 5% = -2.923

Critical value 10% = -2.599

=> P-Value = 0.293. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Life Expectancy"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -3.8091

No. Lags Chosen = 5

Critical value 1% = -3.566

Critical value 5% = -2.92

Critical value 10% = -2.598

=> P-Value = 0.0028. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Birth Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.4883

No. Lags Chosen = 0

Critical value 1% = -3.553

Critical value 5% = -2.915

Critical value 10% = -2.595

=> P-Value = 0.1184. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Death Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.9105

No. Lags Chosen = 10

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.0441. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Infant Mortality Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.986

No. Lags Chosen = 5

Critical value 1% = -3.566

Critical value 5% = -2.92

Critical value 10% = -2.598

=> P-Value = 0.0362. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Fertility Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -1.0891

No. Lags Chosen = 10

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.7194. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

综上所述，根据 ADF 检验的结果，"Life Expectancy"、"Death Rate"和"Infant Mortality Rate"变量是平稳的，而"Population Density"、"Birth Rate"和"Fertility Rate"变量是非平稳的。

我们第三次差分操作，结果如下所示：

Augmented Dickey-Fuller Test on "Population Density"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.919

No. Lags Chosen = 6

Critical value 1% = -3.571

Critical value 5% = -2.923

Critical value 10% = -2.599

=> P-Value = 0.0432. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Life Expectancy"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -6.268

No. Lags Chosen = 4

Critical value 1% = -3.566

Critical value 5% = -2.92

Critical value 10% = -2.598

=> P-Value = 0.0. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Birth Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -4.1035

No. Lags Chosen = 9

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.001. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Death Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -2.5677

No. Lags Chosen = 9

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.0999. Weak evidence to reject the Null Hypothesis.

=> Series is Non-Stationary.

Augmented Dickey-Fuller Test on "Infant Mortality Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -6.7704

No. Lags Chosen = 4

Critical value 1% = -3.566

Critical value 5% = -2.92

Critical value 10% = -2.598

=> P-Value = 0.0. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Fertility Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -4.4434

No. Lags Chosen = 9

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.0002. Rejecting Null Hypothesis.

=> Series is Stationary.

根据更新后的 ADF 检验结果, "Population Density"、"Life Expectancy"、"Birth Rate"、"Infant Mortality Rate"和"Fertility Rate"变量都是平稳的, 而"Death Rate"变量是非平稳的。

我们继续进行第四次差分操作, 得到的结果为:

Augmented Dickey-Fuller Test on "Population Density"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05
Test Statistic = -4.3378
No. Lags Chosen = 5
Critical value 1% = -3.571
Critical value 5% = -2.923
Critical value 10% = -2.599

=> P-Value = 0.0004. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Life Expectancy"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05
Test Statistic = -4.2616
No. Lags Chosen = 3
Critical value 1% = -3.566
Critical value 5% = -2.92
Critical value 10% = -2.598

=> P-Value = 0.0005. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Birth Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05
Test Statistic = -3.4048
No. Lags Chosen = 11
Critical value 1% = -3.593
Critical value 5% = -2.932
Critical value 10% = -2.604

=> P-Value = 0.0108. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Death Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -6.446

No. Lags Chosen = 8

Critical value 1% = -3.581

Critical value 5% = -2.927

Critical value 10% = -2.602

=> P-Value = 0.0. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Infant Mortality Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -4.0283

No. Lags Chosen = 3

Critical value 1% = -3.566

Critical value 5% = -2.92

Critical value 10% = -2.598

=> P-Value = 0.0013. Rejecting Null Hypothesis.

=> Series is Stationary.

Augmented Dickey-Fuller Test on "Fertility Rate"

Null Hypothesis: Data has unit root. Non-Stationary.

Significance Level = 0.05

Test Statistic = -3.3566

No. Lags Chosen = 11

Critical value 1% = -3.593

Critical value 5% = -2.932

Critical value 10% = -2.604

=> P-Value = 0.0125. Rejecting Null Hypothesis.

=> Series is Stationary.

根据更新后的 ADF 检验结果, "Population Density"、"Life Expectancy"、"Birth Rate"、"Death Rate"、"Infant Mortality Rate"和"Fertility Rate"变量都是平稳的。

4.5 选择模型阶数并训练

使用 VAR 模型对差分后的数据进行拟合, 并打印出不同滞后阶数下的模型评估指标。可以使用以下指标来评估模型的拟合优度和复杂度:

AIC (Akaike Information Criterion): AIC 是一种模型选择准则, 它通过考虑拟合优度和模型复杂度来评估模型。AIC 值越小表示模型的拟合优度越好, 同时考虑到了模型的复杂度。AIC 的计算公式为:

$$AIC = -2 \ln L + 2k$$

其中 L 是模型的似然函数值, k 是模型的参数数量。

BIC (Bayesian Information Criterion): BIC 是另一种模型选择准则, 它在考虑拟合优度和模型复杂度时比 AIC 更加严格。与 AIC 类似, BIC 值越小表示模型的拟合优度越好, 但对模型的复杂度有更强的惩罚。BIC 的计算公式为:

$$BIC = -2 \ln L + k \ln n$$

其中 n 是样本数量。

FPE (Final Prediction Error): FPE 是对模型拟合优度的一种衡量, 它用于估计模型在未来预测时的误差。FPE 值越小表示模型的拟合优度越好。FPE 的计算公式为:

$$FPE = (n + k) \times SSE / (n - k)$$

其中 n 是样本数量, k 是模型的参数数量, SSE 是模型的残差平方和。

HQIC (Hannan-Quinn Information Criterion): HQIC 是模型选择准则之一, 类似于 AIC 和 BIC, 它在评估拟合优度和模型复杂度时提供了一种平衡。HQIC 值越小表示模型的拟合优度越好, 考虑了模型的复杂度。HQIC 的计算公式为:

$$HQIC = -2 \ln L + 2k \ln(\ln n)$$

其中 L 是模型的似然函数值, k 是模型的参数数量, n 是样本数量。

这些指标都是用来比较不同滞后阶数下 VAR 模型的拟合优度和模型复杂度。

一般来说，选择具有较小 AIC、BIC、FPE 和 HQIC 值的滞后阶数，可以得到较好的模型拟合效果。但最终选择的滞后阶数可能还需要考虑其他因素，例如模型的解释性、稳定性和经验判断。

检查不同的滞后阶数并计算信息准则的值，以确定 VAR 模型的最优滞后阶数。结果如下所示：

Lag Order = 1

AIC : -23.117769526524256

BIC : -21.570781934752045

FPE : 9.202438613669279e-11

HQIC: -22.5211570787672

Lag Order = 2

AIC : -24.506494199549213

BIC : -21.606819307906466

FPE : 2.4186204713010765e-11

HQIC: -23.391419086156127

Lag Order = 3

AIC : -25.287236190823016

BIC : -21.009509577009887

FPE : 1.288679874056183e-11

HQIC: -23.647256767394826

Lag Order = 4

AIC : -27.59255337958725

BIC : -21.910713283339234

FPE : 1.8074836127427066e-12

HQIC: -25.42135422896446

Lag Order = 5

AIC : -31.321472325054753

BIC : -24.20874674486205

FPE : 8.800249117566737e-14

HQIC: -28.61290585636472

Lag Order = 6

AIC : -35.41296861736406

BIC : -26.841864409597548

FPE : 6.5615520407066e-15

HQIC: -32.16110363643813

结果总结如下图所示：

VAR Order Selection (* highlights the minimums)				
	AIC	BIC	FPE	HQIC
0	-24.35	-24.12	2.660e-11	-24.26
1	-25.20	-23.58	1.149e-11	-24.59
2	-24.83	-21.82	1.779e-11	-23.69
3	-25.25	-20.85	1.398e-11	-23.58
4	-27.40	-21.61	2.363e-12	-25.21
5	-31.59	-24.40	7.450e-14	-28.86
6	-35.41*	-26.84*	6.562e-15*	-32.16*

图 7：不同的滞后阶数并计算信息准则的值

根据提供的 VAR Order Selection 摘要结果，可以进行以下分析：

AIC：最小值出现在滞后阶数为 6 的情况下，为-35.41。较低的 AIC 值表示模型的拟合度和简洁度较好。

BIC：最小值同样出现在滞后阶数为 6 的情况下，为-26.84。较低的 BIC 值也表示模型的拟合度和简洁度较好。

FPE：最小值同样出现在滞后阶数为 6 的情况下，为 6.562e-15。较小的 FPE 值表示模型的预测能力较好。

HQIC：最小值同样出现在滞后阶数为 6 的情况下，为-32.16。较低的 HQIC 值表示模型的拟合度和简洁度较好。

综合考虑这些信息准则，滞后阶数为 6 在各个准则下都显示出最优的结果。因此，可以选择滞后阶数为 6 的 VAR 模型作为最佳模型。

4.6 训练选定的 VAR 模型 (p)

我们选用上面选择的滞后阶数为 6 的 VAR 模型作为最终训练的模型，结果如下所示：

Summary of Regression Results			
=====			
Model:	VAR		
Method:	OLS		
Date:	Fri, 26, May, 2023		
Time:	15:16:49		

No. of Equations:	6.00000	BIC:	-26.8419
Nobs:	49.0000	HQIC:	-32.1611
Log likelihood:	672.450	FPE:	6.56155e-15
AIC:	-35.4130	Det(Omega_mle):	2.24488e-16

图 8：模型总结的结果

Results for equation Population Density:

	coefficient	std. error	t-stat	prob

const	0.001058	0.007418	0.143	0.887
L1.Population Density	-0.019182	0.236940	-0.081	0.935
L1.Life Expectancy	1.506466	0.962088	1.566	0.117
L1.Birth Rate	0.279625	0.205528	1.361	0.174
L1.Death Rate	2.479600	1.580958	1.568	0.117
L1.Infant Mortality Rate	-0.043617	0.055588	-0.785	0.433
L1.Fertility Rate	-2.616888	1.733397	-1.510	0.131
L2.Population Density	-0.121801	0.170690	-0.714	0.475
L2.Life Expectancy	0.432195	0.450579	0.959	0.337
L2.Birth Rate	0.236517	0.102113	2.316	0.021
L2.Death Rate	0.077235	0.567304	0.136	0.892
L2.Infant Mortality Rate	0.060637	0.055493	1.093	0.275
L2.Fertility Rate	-1.250691	0.704138	-1.776	0.076
L3.Population Density	0.178620	0.201948	0.884	0.376
L3.Life Expectancy	-0.336573	0.455533	-0.739	0.460
L3.Birth Rate	0.468890	0.120406	3.894	0.000
L3.Death Rate	-0.936001	0.527509	-1.774	0.076
L3.Infant Mortality Rate	0.072342	0.059295	1.220	0.222
L3.Fertility Rate	-2.494149	0.807052	-3.090	0.002
L4.Population Density	0.047931	0.164100	0.292	0.770
L4.Life Expectancy	-0.076993	0.433851	-0.177	0.859
L4.Birth Rate	0.227109	0.136509	1.664	0.096
L4.Death Rate	-0.769848	0.501752	-1.534	0.125
L4.Infant Mortality Rate	0.077409	0.068588	1.129	0.259
L4.Fertility Rate	-1.585850	0.841425	-1.885	0.059
L5.Population Density	0.195027	0.137900	1.414	0.157
L5.Life Expectancy	-0.510172	0.372549	-1.369	0.171
L5.Birth Rate	0.300801	0.125220	2.402	0.016
L5.Death Rate	-0.601713	0.402332	-1.496	0.135
L5.Infant Mortality Rate	-0.080477	0.062443	-1.289	0.197
L5.Fertility Rate	-2.220717	0.782166	-2.839	0.005
L6.Population Density	-0.356713	0.145543	-2.451	0.014
L6.Life Expectancy	-1.478544	0.702114	-2.106	0.035
L6.Birth Rate	0.311778	0.235749	1.323	0.186
L6.Death Rate	-1.756203	1.139692	-1.541	0.123
L6.Infant Mortality Rate	0.003958	0.095916	0.041	0.967
L6.Fertility Rate	-1.084261	1.111179	-0.976	0.329
=====				

图 9：Results for equation Population Density

其他变量的结果这里不再展示，详情请见附录内容。

最后展示一下残差的相关系数矩阵，如下图所示：

Correlation matrix of residuals						
	Population Density	Life Expectancy	Birth Rate	Death Rate	Infant Mortality Rate	Fertility Rat
Population Density	1.000000	-0.554084	-0.013528	0.624015	0.499786	0.16362
Life Expectancy	-0.554084	1.000000	-0.271801	-0.987274	-0.910613	-0.31894
Birth Rate	-0.013528	-0.271801	1.000000	0.309979	0.423525	0.95393
Death Rate	0.624015	-0.987274	0.309979	1.000000	0.923034	0.39624
Infant Mortality Rate	0.499786	-0.910613	0.423525	0.923034	1.000000	0.45705
Fertility Rate	0.163625	-0.318947	0.953936	0.396243	0.457057	1.00000

图 10：残差的相关系数矩阵

总体而言，除了常数项之外，其他系数没有在统计上表现出显著性。这意味着在滞后阶数为 6 的 VAR 模型中，常数项对于解释方程的变动有限影响。

4.7 检查残差的序列相关性

残差的序列相关性用于检查残差中是否存在任何剩余模式（误差）。这对我们意味着什么？

如果残差中还存在任何相关性，那么时间序列中仍有一些模式有待模型解释。在这种情况下，典型的行动过程是增加模型的顺序或将更多预测因子引入系统，或者寻找不同的算法来对时间序列进行建模。

因此，检查序列相关性是为了确保模型能够充分解释时间序列中的方差和模式。

要检查 VAR 模型的残差序列相关性，可以使用 Durbin-Watson（DW）统计量。DW 统计量用于检验残差序列中是否存在自相关性。

$$DW = \frac{\sum_{t=2}^T ((e_t - e_{t-1})^2)}{\sum_{t=1}^T e_t^2}$$

DW 统计量的取值范围为 0 到 4，其值越接近 2，表示残差序列的自相关性越低（接近于无自相关性）。一般来说，DW 统计量在 1.5 到 2.5 之间被认为是较理想的范围。

我们得出的结果为：

Population Density : 2.14

Life Expectancy : 1.95

Birth Rate : 2.03

Death Rate : 1.96

Infant Mortality Rate : 1.91

Fertility Rate : 1.94

DW 统计量为 2.14，接近于 2，表明残差序列在该变量上没有显著的自相关性。这是一个好的特征，表明 VAR 模型在对这些变量进行建模时能够捕捉到数据中的大部分信息，并且模型的残差序列不包含明显的自相关性。

4.8 结果预测

为了进行预测，VAR 模型期望过去数据的观测值达到滞后顺序数。这是因为，VAR 模型中的项本质上是数据集中各种时间序列的滞后，因此我们需要为其提供模型使用的滞后顺序所指示的尽可能多的先前值。

根据模型拟合的结果，滞后阶数（lag order）为 6。这意味着在进行预测时，需要提供最后 6 个观测值的差分数据作为输入。

让我们预测一下：

	Population Density_4d	Life Expectancy_4d	Birth Rate_4d	Death Rate_4d	Infant Mortality Rate_4d	Fertility Rate_4d
Year						
2018	-0.014230	0.028869	0.576380	0.033933	0.403380	0.086829
2019	-0.099357	-0.002861	-0.913989	-0.065179	-0.414440	-0.133082
2020	0.112947	-0.060190	0.409970	0.075628	0.340612	0.063655
2021	0.069413	0.089148	-0.135240	-0.075073	-0.433698	-0.027194

图 11：初始的预测结果

预测是生成的，但它是模型使用的训练数据的规模。因此，要使其恢复到原始比例，我们需要将其去差分，就像对原始输入数据进行差分的次数一样。经过四次反转之后，我们的预测结果如下所示：

	Population Density_forecast	Life Expectancy_forecast	Birth Rate_forecast	Death Rate_forecast	Infant Mortality Rate_forecast	Fertility Rate_forecast
Year						
2018	147.255770	76.238869	12.107380	7.353933	9.896380	1.707829
2019	148.012183	76.554876	11.614770	7.345688	9.408320	1.671575
2020	148.881542	76.810693	11.532131	7.413070	9.260871	1.698977
2021	149.820316	77.155658	11.314253	7.405379	8.679725	1.699185

图 12：四次反转后的预测结果

预测恢复到原来的规模。让我们根据测试数据的实际值绘制预测值。

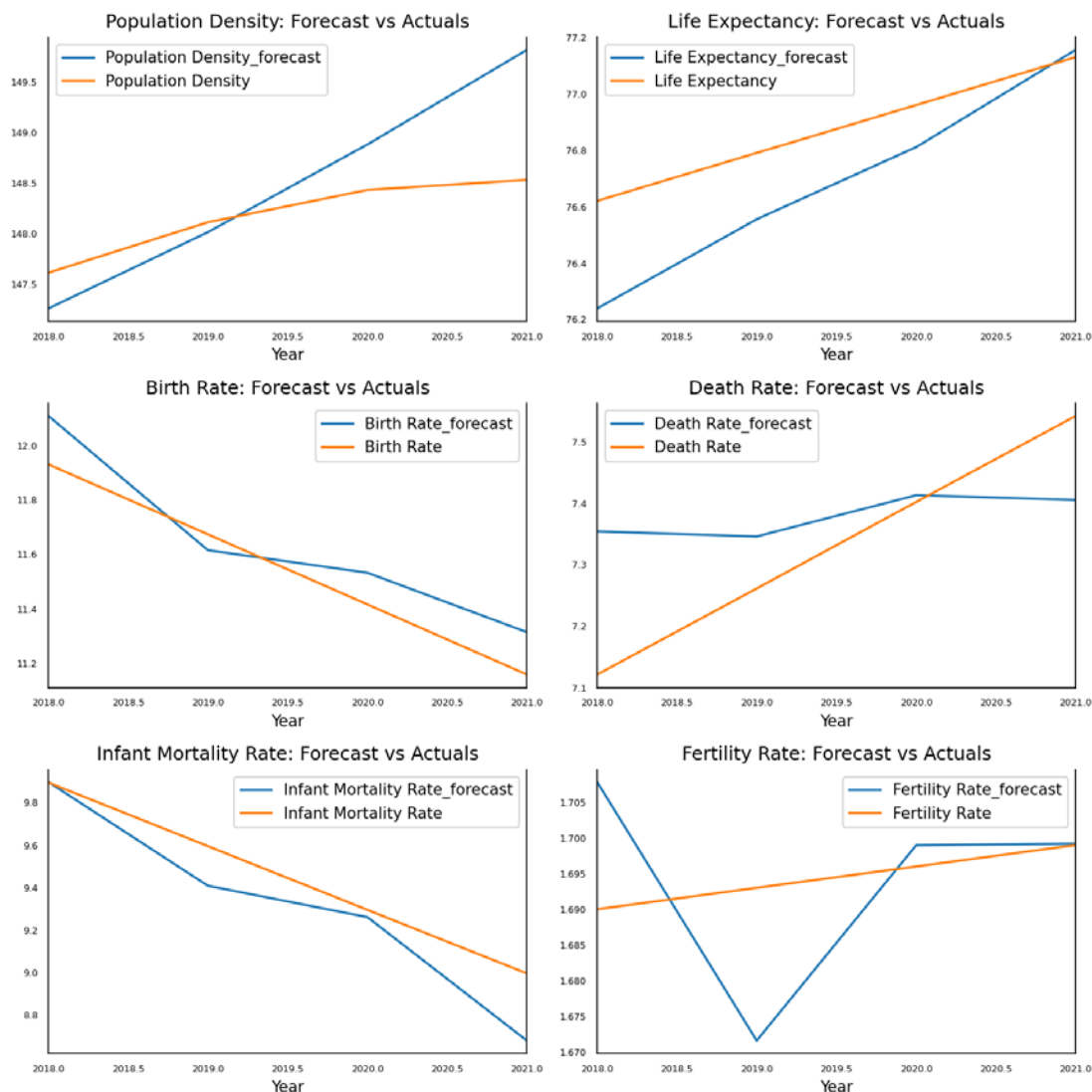


图 13: 预测结果和真实结果的可视化

4.9 评估预测

为了评估预测，让我们计算一组全面的指标，即 MAPE, ME, MAE, MPE, RMSE, corr 和 minmax。

我们简单介绍一下各个指标及其作用：

平均绝对百分比误差 (Mean Absolute Percentage Error, MAPE)，计算实际值与预测值之间的百分比误差的平均值。平均误差 (Mean Error, ME)，计算实际值与预测值之间的平均差。平均绝对误差 (Mean Absolute Error, MAE)，计算实际值与预测值之间的平均绝对差。平均百分比误差 (Mean Percentage Error, MPE)，计算实际值与预测值之间的百分比误差的平均值。均方根误差 (Root Mean Square Error, RMSE)，计算实际值与预测值之间的均方根差。相关系数 (Correlation

Coefficient), 计算实际值与预测值之间的相关性。最大最小误差 (Min-Max Error), 计算实际值与预测值之间的最大最小差。

表 1: 各变量的评价指标值

	MAPE	ME	MAE	MPE	RMSE	CORR	MINMAX
Population Density	0.0037	0.3225	0.5485	0.0022	0.7078	0.946	0.0037
Life Expectancy	0.0026	-0.185	0.1978	-0.002	0.2364	0.9985	0.0026
Birth Rate	0.011	0.0976	0.1267	0.0085	0.1346	0.9486	0.0109
Death Rate	0.016	0.048	0.1163	0.0069	0.1416	0.8278	0.0157
Infant Mortality Rate	0.0146	-0.134	0.1349	-0.015	0.1844	0.9783	0.0146
Fertility Rate	0.0063	-0.000	0.0106	-0.000	0.014	0.0121	0.0062

可以对每个指标进行以下分析:

- 平均绝对百分比误差(MAPE): 所有指标的 MAPE 值都在较低的范围內, 表明预测与实际值之间的百分比误差相对较小。
- 平均误差 (ME): 大部分指标的 ME 接近零, 表示平均预测误差接近于零, 但人口密度的 ME 稍高。
- 平均绝对误差 (MAE): 所有指标的 MAE 值都在相对较小的范围内, 表示预测值与实际值之间的平均绝对差相对较小。
- 平均百分比误差 (MPE): 所有指标的 MPE 值接近零, 表示预测值与实际值之间的百分比误差相对较小。
- 均方根误差 (RMSE): RMSE 值较小, 表明预测值与实际值之间的差异相对较小。
- 相关系数 (CORR): 大部分指标的相关系数接近于 1, 表示预测值与实际值之间存在较强的正相关性。
- 最大最小误差 (MINMAX): 最大最小误差值较小, 表示预测值与实际值之间的最大最小差相对较小。

综合来看, 这些指标结果表明预测模型的准确性较高, 预测值与实际值之间的差异相对较小, 并且存在较强的相关性。然而, 对于人口密度和出生率指标, 可能存在稍微较高的预测误差和差异。

五、 总结

根据采用 VAR 模型对中国人口数据进行预测的研究，本文旨在探索人口变量之间的关联关系，并利用该模型对未来人口发展进行预测。总体而言，结果表明预测模型在大多数指标上表现出较高的准确性，预测值与实际值之间的差异相对较小，并且变量之间存在较强的相关性。

通过 VAR 模型的预测结果，我们能够观察到人口变量的趋势和演变情况。预测结果显示出生口密度、出生率等指标可能存在稍微较高的预测误差和差异，这可能源于该模型对于人口变量之间复杂关系的建模挑战。然而，综合考虑其他指标的预测结果，我们可以认为整体预测模型具有较高的可靠性和准确性。

此外，研究结果也揭示了人口变量之间的相互作用和影响因素。通过探索 VAR 模型的相关系数，我们可以进一步理解不同变量之间的关联性，从而对人口发展的驱动因素进行分析和解释。这对于制定人口政策和规划具有重要的参考意义。

然而，本研究也存在一些局限性。首先，模型的预测能力受到数据质量和可用性的限制。其次，模型的假设和参数选择可能对预测结果产生一定的影响。进一步的研究可以考虑引入更多的影响因素和扩展模型，以提高预测的准确性和稳定性。

综上所述，本研究采用 VAR 模型对中国人口数据进行预测，并取得了较好的预测结果。预测模型具有较高的准确性和可靠性，并提供了对人口变量之间关系和趋势的洞察。然而，需要在实际应用中进一步考虑模型的局限性，并结合其他方法和数据进行综合分析，以支持人口政策的制定和人口发展的规划。

六、 参考文献

- [1]陈六傅, 刘厚俊. 人民币汇率的价格传递效应—基于 VAR 模型的实证分析[J]. 金融研究, 2007(04A):13.
- [2]周应恒, 邹林刚. 中国大豆期货市场与国际大豆期货市场价格关系研究—基于 VAR 模型的实证分析[J]. 农业技术经济, 2007(1):8.
- [3]王来福, 郭峰. 货币政策对房地产价格的动态影响研究—基于 VAR 模型的实证[J]. 财经问题研究, 2007(11):5.
- [4]张晶. 中国货币政策区域效应差异及其原因研究—结构 VAR 模型下的实证分析[J]. 广东金融学院学报, 2006.
- [5]段成荣, 杨书章, 高书国. 21 世纪上半叶我国各级学校适龄人口数量变动趋势分析[J]. 人口与经济, 2000(4):8.
- [6] Statsmodels 官方文档: https://www.statsmodels.org/stable/vector_ar.html
- [7]<https://www.machinelearningplus.com/time-series/vector-autoregression-examples-python/>