



پروژه‌ی دوم  
درس تحلیل رگرسیون

## واکاوی الگوهای پنهان فقر در تهران

اعضای گروه:

ارشیا دهقان (۴۰۱۱۰۰۳۸۲)

ابوالفضل مسلمی (۴۰۱۱۰۰۵۰۶)

علی محمدزاده شبستری (۴۰۱۱۰۶۴۸۲)

استاد:

جناب آقای علیرضا کدیور

۱۴۰۴ مرداد

## فهرست مطالب

۱.	چکیده.....	۳
۲.	مقدمه.....	۳
۳.	رویکرد مدل‌سازی.....	۳
۴.	روش‌شناسی.....	۴
۴.	۴.۱. منبع داده‌ها.....	۴
۴.	۴.۲. متغیر وابسته.....	۴
۴.	۴.۳. متغیرهای مستقل (ویژگی‌ها).....	۴
۵.	۴.۴. پیش‌پردازش داده‌ها.....	۵
۵.	۴.۵. استراتژی اول: مدل رگرسیون لاجستیک.....	۵
۶.	۴.۶. استراتژی دوم: مدل جمعی تعمیم‌یافته‌ی نیمه‌پارامتریک (GAM).....	۶
۷.	۵. بحث درباره‌ی نتایج.....	۷
۷.	۵.۱. مدل رگرسیون لجستیک.....	۷
۹.	۵.۲. مدل جمعی تعمیم‌یافته.....	۹
۱۳.	۶. محدودیت‌ها و مسیر آینده.....	۱۳
۱۴.	۷. نتیجه‌گیری و فرصت‌های آینده.....	۱۴
۱۴.	۸. پیوست.....	۱۴

## ۱. چکیده

این پژوهش با هدف شناسایی و مدل‌سازی دقیق عوامل کلیدی اجتماعی، جمعیت‌شناختی و اقتصادی مرتبط با فقر نسبی در میان خانوارهای تهرانی انجام شده است. با استفاده از داده‌های طرح هزینه و درآمد خانوار مرکز آمار ایران در سال‌های ۱۳۹۹ تا ۱۴۰۱ (۵۸۲۳ خانوار)، یک استراتژی تحلیلی دو مرحله‌ای طراحی شد. در مرحله‌ی نخست، مدل رگرسیون لجستیک نقش برجسته‌ی تحصیلات، تملک و سن سرپرست خانوار را تأیید کرد، اما در آزمون نیکویی برازش<sup>۱</sup> عملکرد ضعیفی داشت که نشان‌دهنده‌ی ناکارآمدی آن در مدل‌سازی پیچیدگی‌های پدیده‌ی فقر است. در گام دوم، یک مدل جمیع تعیین‌بافته‌ی نیمه‌پارامتریک<sup>۲</sup> (GAM) برای تحلیل دقیق‌تر برازش شد. نتایج مدل GAM علاوه بر تأیید نتایج مدل قبلی، وجود رابطه‌ای غیرخطی U-شکل بین سن سرپرست خانوار و احتمال فقر، و نیز الگوهای پیچیده بین سهم انواع درآمدها، نسبت واپستگی<sup>۳</sup> و فقر را آشکار ساخت. اگرچه متغیرهای جنسیت و وضعیت تأهل سرپرست در این مدل جامع معنادار نبودند، این به معنای بی‌اهمیتی آنها نیست. بلکه نشان می‌دهد تأثیرشان از طریق کانال‌های اقتصادی مانند اشتغال و تحصیلات عمل می‌کند. بنابراین، تحلیل‌های آتی باید بر واکاوی جداگانه‌ی این عوامل در میان گروه‌های آسیب‌پذیر، مشخصاً خانوارهای زن سرپرست، متمرکز شوند.

## ۲. مقدمه

فقر همچنان یکی از چالش‌های چندلایه و پایدار در جهان امروز برای سیاست‌گذاران و پژوهشگران به شمار می‌رود. در ایران، شهر تهران به عنوان پایتخت اقتصادی، نمونه‌ای فشرده از روندهای اقتصادی، دگرگونی‌های جمعیتی و نابرابری‌های اجتماعی در کشور را به تصویر می‌کشد. این شهر در کنار فرصت‌های فراوان، با آسیب‌پذیری‌ها و نابرابری‌های جدی خانوارها روبروست. درک دقیق عواملی که خانوارهای فقیر را از غیرفقیر تمایز می‌کنند، برای طراحی سیاست‌های اجتماعی هدفمند و کارآمد ضروری است. این پژوهش با بهره‌گیری از مدل‌های آماری بر روی داده‌های سه سال اخیر، به دنبال شناسایی دقیق و کمی‌سازی تأثیر این عوامل، به ویژه روابط غیرخطی میان آن‌هاست.

## ۳. رویکرد مدل‌سازی

برخلاف بسیاری از مدل‌های آماری که تمرکز آن‌ها بر دقت پیش‌بینی است، هدف اصلی این مطالعه تفسیر روابط ساختاری فقر است. سوال اصلی تحقیق آن است که: «کدام ویژگی‌های اجتماعی، جمعیت‌شناختی و اقتصادی، بیشترین ارتباط را با فقر نسبی در خانوارهای تهرانی دارند و شکل دقیق این روابط چگونه است؟» مدل‌های پارامتری متدائل مانند رگرسیون لجستیک، در عین قابل تفسیر بودن، بر فرض‌های محدود کننده‌ای مانند خطی بودن روابط (در مقیاس log-odds) تکیه دارند. این فرض‌ها ممکن است شکل واقعی و پیچیده‌ی عوامل مؤثر بر فقر را پنهان یا دگرگون سازند. برای غلبه بر محدودیت‌های مذکور، از یک رویکرد دو مرحله‌ای استفاده شده است: ابتدا مدل رگرسیون لجستیک برای برآورد پایه به کار رفته و سپس مدل GAM برای کشف روابط غیرخطی به کار گرفته شده است. این رویکرد امکان تحلیل دقیق‌تر و عمیق‌تری از عوامل مرتبط با فقر در تهران را فراهم می‌کند.

۱ goodness of fit  
Generalized Additive Model ۲

۳ نسبت واپستگی (Dependency Ratio) یک شاخص جمعیتی است که نشان‌دهنده نسبت افراد غیرفعال به افراد فعال است. این نسبت برای ارزیابی بار اقتصادی جمعیت غیرفعال بر جمعیت فعال استفاده می‌شود و می‌تواند تأثیرات مهمی بر اقتصاد یک کشور داشته باشد.

## ۴. روش‌شناسی

### ۱.۴. منبع داده‌ها

این مطالعه از داده‌های طرح سالانه‌ی هزینه و درآمد خانوار (HEIS) که توسط مرکز آمار ایران در سال‌های ۱۳۹۹ تا ۱۴۰۱ اجرا شده، استفاده می‌کند. داده‌ها از طریق کتابخانه‌ی HBSIR در پایتون استخراج شده‌اند. جامعه‌ی مورد بررسی به خانوارهای ساکن استان تهران محدود شده است. پس از حذف مشاهدات دارای داده‌ی گمشده و رفع مشکلات مربوط به ناپایداری مدل (که در بخش‌های بعدی توضیح داده شده است)، نمونه‌ی نهایی شامل ۵۸۲۳ خانوار است. واحد تحلیل در این مطالعه، خانوار بوده و در تمامی تحلیل‌ها از وزن‌های ارائه‌شده‌ی رسمی استفاده شده است تا نتایج قابلیت عمیمی به جمعیت استان تهران را داشته باشد.

### ۲.۴. متغیر وابسته

متغیر وابسته‌ی این مطالعه، با عنوان فقر نسبی (Is\_Poor)، به صورت یک متغیر دودویی لحاظ شده و مبتنی بر تعریف استاندارد فقر نسبی در مطالعات اقتصادی تنظیم شده است:

- **هزینه‌ی خالص تعدیل شده:** هزینه‌ی خالص سالانه‌ی هر خانوار بر اندازه‌ی معادل آن تقسیم شده است. برای محاسبه‌ی اندازه‌ی معادل خانوار از مقیاس OECD تعدیل شده<sup>۴</sup> استفاده شده تا تفاوت در ترکیب خانوارها و صرفهجویی‌های ناشی از مقیاس در نظر گرفته شود.
- **خط فقر سالانه:** برای هر یک از سال‌های ۱۳۹۹ تا ۱۴۰۱، خط فقر نسبی به صورت ۶۰٪ میانه‌ی وزنی هزینه‌ی خالص تعدیل شده‌ی خانوارهای ساکن تهران در همان سال تعریف شده است. این تعریف، فقر را به نسبت استانداردهای زندگی همان دوره در نظر می‌گیرد.
- **طبقه‌بندی خانوارها:** خانوارهایی که هزینه‌ی خالص تعدیل شده‌ی آنها کمتر از خط فقر سال مربوطه باشد، به عنوان خانوار فقیر ( $Is\_Poor = 1$ ) و سایر خانوارها به عنوان غیرفقیر ( $Is\_Poor = 0$ ) کدگذاری شده‌اند.

### ۳.۴. متغیرهای مستقل (ویژگی‌ها)

با استناد به ادبیات اقتصادی معتبر، مجموعه‌ای از متغیرها برای پیش‌بینی فقر نسبی انتخاب شده‌اند:

- **ویژگی‌های جمعیت‌شناختی:**
  - سن سرپرست خانوار (Head\_Age): متغیر طبقه‌ای (بازه‌بندی شده)
  - جنسیت سرپرست خانوار (Head\_Sex): متغیر طبقه‌ای دودویی
  - سطح تحصیلات سرپرست خانوار (Head\_Education\_Level): متغیر طبقه‌ای
  - وضعیت فعالیت اقتصادی سرپرست خانوار (Head\_Activity\_Status): متغیر طبقه‌ای
  - وضعیت تأهل سرپرست خانوار (Head\_Marital\_Status): متغیر طبقه‌ای

<sup>۴</sup> مقیاس OECD-Modified به بزرگترین فرد وزن ۱، به افراد بالای ۱۴ سال وزن ۰/۵ و به هر فرد زیر ۱۴ سال وزن ۰/۳ می‌دهد تا اندازه‌ی معادل خانوار را تخمین بزند.

#### • ساختار خانوار:

- اندازه‌ی خانوار (Household\_Size): متغیر طبقه‌ای (بازه‌بندی شده)
- نسبت وابستگی شغلی (Dependency\_Ratio\_Employment): متغیر پیوسته
- سکونت در شهر یا روستا (Urban\_Rural): متغیر طبقه‌ای دودویی
- نحوه‌ی تصرف مسکن (Tenure\_Type): متغیر طبقه‌ای

#### • منبع درآمد خانوار (متغیرهای پیوسته بین صفر و یک):

- سهم درآمد از کمک و یارانه (Aid\_Income\_Share)
- سهم درآمد از کار و دستمزد (Labor\_Income\_Share)
- سهم درآمد از حقوق بازنشستگی (Pension\_Income\_Share)
- سهم درآمد از سرمایه (Capital\_Income\_Share)
- سهم درآمد از اجاره‌ی ضمنی (Imputed\_Rent\_Share)
- سهم درآمد از سایر منابع غیرنقدی (Other\_Noncash\_Income\_Share)

### ۴.۴. پیش‌پردازش داده‌ها

ویژگی‌های معرفی شده طی یک فرآیند انتخاب و پالایش مکرر مورد ارزیابی قرار گرفتند. در این فرایند، متغیرها و دسته‌بندی‌های آن‌ها تعديل شدند تا ترکیب بهینه برای مدل حاصل شود. از جمله شاخص‌های به‌کاررفته در این تصمیم‌گیری، عامل تورم واریانس (VIF) بود که در مواردی مقادیر بالای ۱۰ نشان‌دهنده‌ی هم‌خطی و نامناسب‌بودن برخی متغیرها بود. هم‌چنین در یک فرآیند رفت و برگشتی، پس از هر بار اجرای مدل رگرسیون لاجستیک تغییرات کوچکی در پیش‌پردازش داده‌ها ایجاد کردیم تا در نهایت به حالتی متعادل و مناسب برسد.

در راستای بهبود همگرایی مدل لجستیک و امکان مقایسه‌ی معنادار ضرایب، متغیرهای پیوسته پیش از مدل‌سازی استانداردسازی شده‌اند.

❖ جزئیات مربوط به پیش‌پردازش داده‌ها در پیوست ۱، و مراحل آن در نوت‌بوک مربوطه آمده‌اند.

### ۴.۵. استراتژی اول: مدل رگرسیون لاجستیک

استراتژی مدل‌سازی ما در دو گام اصلی طراحی شد. در گام نخست، برای ایجاد یک مدل پایه و درک اولیه از عوامل تعیین‌کننده فقر، یک مدل رگرسیون لجستیک وزن‌دار برآورش شد. انتخاب این مدل به دلیل قابلیت بالای آن در تفسیر ضرایب به صورت نسبت شانس<sup>۵</sup> و همچنین محاسبه‌ی اثرات حاشیه‌ای<sup>۶</sup> بود که به ما اجازه می‌داد تأثیر هر متغیر بر احتمال فقر را به صورت کمی بیان کنیم. وزن‌های خانوار نیز در این مدل لحاظ گردید تا نتایج به کل جامعه مورد مطالعه قابل تعمیم باشد.

در فرآیند برآزش این مدل، با چالش‌های متعددی روبرو شدیم. یکی از مهم‌ترین آن‌ها، پدیده‌ی جدایش تقریباً کامل<sup>۷</sup> در داده‌ها بود؛ حالتی که در آن برخی سطوح از متغیرهای پیش‌بینی‌کننده (مانند داشتن تحصیلات دکترا) به طور تقریباً کامل نتیجه‌ی وابسته (فقیر نبودن) را پیش‌بینی می‌کردند. این پدیده منجر به ضرایب بسیار بزرگ و ناپایدار در مدل می‌شد که با ایجاد تغییراتی در پیش‌پردازش داده‌ها و تغییر دسته‌بندی‌ها رفع شد.

علاوه بر این، برای آزمودن روابط غیرخطی بالقوه با استناد به متون علمی، چندین مشخصه‌سازی مختلف را امتحان کردیم. به عنوان مثال، توان دوم سن سرپرست خانوار به مدل اضافه شد. همچنین، اثرات متقابل<sup>۸</sup> بین متغیرهای کلیدی مانند سطح تحصیلات و جنسیت سرپرست مورد بررسی قرار گرفت تا مشخص شود آیا تأثیر یک متغیر در سطوح مختلف متغیر دیگر، متفاوت است یا خیر. با وجود این تلاش‌ها، همان‌طور که در ادامه توضیح داده خواهد شد، محدودیت‌های ذاتی مدل لجستیک در شناسایی الگوهای پیچیده، ما را به سمت استفاده از مدل‌های انعطاف‌پذیرتر سوق داد.

## ۴.۶. استراتژی دوم: مدل جمعی تعمیم‌یافته‌ی نیمه‌پارامتریک (GAM)

اگرچه مدل نهایی رگرسیون لجستیک از نظر آماری پایدار بود، اما در آزمون نیکوبی برآزش هاسمر-لمشو مردود شد که نشان‌دهنده اختلاف معنادار بین پیش‌بینی‌های خطی مدل و واقعیت مشاهده شده بود. این نتیجه حاکی از آن بود که فرض خطی بودن رابطه‌ی متغیرهای مستقل با لگاریتم نسبت شانس در مدل نقض شده است. برای غلبه بر این محدودیت، به مدل GAM روی آورديم.

مدل GAM با افزودن توابع هموار غیرپارامتریک به مدل خطی تعمیم‌یافته، امکان شناسایی شکل واقعی رابطه‌ی متغیرهای پیوسته با متغیر وابسته را فراهم می‌سازد، به‌طوری که خود داده‌ها فرم تابعی را تعیین می‌کنند. مدل لجستیک GAM استفاده شده در این پژوهش دارای فرم کلی زیر است:

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + f_1(x_1) + f_2(x_2) + \dots + f_k(x_k)$$

در این رابطه،  $p$  احتمال فقر و  $f_i(x_i)$  توابع هموار از متغیرهای مستقل هستند. برای متغیرهای پیوسته از توابع هموارساز اسپلاین (که در پیاده‌سازی با  $\text{S}$  مشخص شده) و برای متغیرهای دسته‌ای از توابع فاکتور (مشخص شده با  $f$ ) استفاده گردید. این روش قابلیت بالایی در شناسایی الگوهای پیچیده و غیرخطی (نظیر روابط آستانه‌ای یا غیر یکنوا) دارد که در مدل‌های خطی قابل شناسایی نیستند. برای جلوگیری از بیش‌برآزش<sup>۹</sup>، میزان همواری توابع اسپلاین از طریق جستجوی شبکه‌ای<sup>۱۰</sup> بر روی پارامتر تنظیم‌کننده ( $\lambda$ ) و با استفاده از معیار اعتبارسنجی متقابل تعمیم‌یافته<sup>۱۱</sup> (GCV) بهینه شد.

Quasi-Complete Separation ۷

Interaction Effects ۸

overfitting ۹

Grid Search ۱۰

Generalized Cross Validation ۱۱

## ۵. بحث درباره نتایج

### ۱.۵. مدل رگرسیون لجستیک

اساس این تحلیل بر مقایسه استوار است؛ بنابراین برای هر متغیر طبقه‌ای (مانند سطح تحصیلات)، یک دسته را به عنوان «مرجع»<sup>۱۲</sup> تعریف کرده‌ایم. مقدار مرجع متغیرها به شرح زیر است:

▪	وضعیت تأهل سرپرست:	متأهل	▪	تحصیلات سرپرست:	دیپلم
▪	وضعیت فعالیت سرپرست:	شاغل	▪	جنسیت سرپرست:	مرد
▪	نحوه تصرف مسکن:	صاحب	▪	سن سرپرست:	۴۰-۵۰
▪			▪	منطقه سکونت:	روستا

نتایج این مدل اولیه، هرچند از نظر عددی دقیق نیستند، اما سرنخ‌های ارزشمندی را آشکار ساختند:

**عوامل کلیدی:** مطابق جدول ۱، مدل تأیید کرد که سرمایه‌ی انسانی (سطح تحصیلات و وضعیت اشتغال سرپرست)، ساختار دارایی (نوع تصرف مسکن و سهم درآمدهای غیرنقدی) و عوامل جمعیت‌شناختی (سن سرپرست و نسبت وابستگی) قوی‌ترین و معنادارترین پیش‌بینی‌کننده‌های ریسک فقر هستند. به طور کلی، تحصیلات بالاتر، زندگی در شهر و مالکیت مسکن به عنوان عوامل محافظتی عمل کرده، در حالی که بیکاری و نسبت وابستگی بالا، ریسک فقر را به شدت افزایش می‌دهند.

**متغیرهای غیرمعنادار:** نکته‌ی قابل توجه، عدم معناداری آماری متغیرهایی مانند جنسیت، وضعیت تأهل و اندازه خانوار در این مدل بود. این یافته نشان می‌دهد که تأثیر این عوامل احتمالاً از طریق متغیرهای اقتصادی مستقیم‌تری که در مدل حضور دارند (مانند تحصیلات، اشتغال و نسبت وابستگی) جذب و نمایندگی شده است.

جدول ۱ – تأثیر متغیرهای معنادار رگرسیون لجستیک (با  $p < 0.05$ ) بر فقر

	Coefficient	Odds_Ratio	Marginal_Effect	p_value
C(Head_Education_Level) [T.Illiterate]	1.7041	5.4963	0.2484	0.0000
C(Tenure_Type) [T.Owner]	1.4962	4.4645	0.2181	0.0000
C(Head_Activity_Status) [T.Other]	1.0465	2.8477	0.1525	0.0000
C(Head_Activity_Status) [T.Unemployed]	1.0399	2.8289	0.1516	0.0002
C(Tenure_Type) [T.Tenant]	0.9251	2.5220	0.1348	0.0000
C(Head_Education_Level) [T.Under_Diploma]	0.8995	2.4583	0.1311	0.0000
C(Head_Activity_Status) [T.Housekeeper]	0.8080	2.2435	0.1178	0.0003
Dependency_Ratio_Employment	0.5168	1.6766	0.0753	0.0000
Capital_Income_Share	-0.1173	0.8893	-0.0171	0.0017
Other_NonCash_Income_Share	-0.2530	0.7764	-0.0369	0.0000
C(Head_Age) [T.40-50]	-0.4381	0.6453	-0.0639	0.0239
C(Head_Age) [T.50-60]	-0.4591	0.6318	-0.0669	0.0194
C(Head_Age) [T.60-70]	-0.6028	0.5473	-0.0879	0.0050
C(Urban_Rural) [T.Urban]	-0.9051	0.4045	-0.1319	0.0000
Imputed_Rent_Share	-1.0583	0.3471	-0.1543	0.0000
C(Head_Education_Level) [T.Bachelor]	-1.1176	0.3271	-0.1629	0.0000
C(Head_Age) [T.Above_70]	-1.1796	0.3074	-0.1719	0.0000
C(Head_Education_Level) [T.Masters_PhD]	-1.8997	0.1496	-0.2769	0.0000

## نکته‌ی اول | عدم معناداری آماری متغیر جنسیت سرپرست خانوار

یکی از یافته‌های قابل تأمل در این مدل، عدم معناداری آماری متغیر «جنسیت سرپرست خانوار» در پیش‌بینی احتمال فقر است ( $p-value = 0.82$ ). در نگاه اول، این نتیجه ممکن است با تصور رایج مبنی بر آسیب‌پذیری بیشتر خانوارهای زن سرپرست در تضاد باشد. با این حال، این یافته به معنای عدم وجود «فقر زنانه» نیست؛ بلکه تفسیری عمیق‌تر را ایجاد می‌کند. در یک مدل رگرسیون چندمتغیره، اثر هر متغیر پس از کنترل سایر متغیرها سنجیده می‌شود. عدم معناداری جنسیت در اینجا نشان می‌دهد که اثر واقعی آن احتمالاً از طریق سایر متغیرهای موجود در مدل (مانند سطح تحصیلات، اشتغال و درآمد) جذب و نمایندگی شده است. به عبارت دیگر، اگر خانوارهای زن سرپرست به طور متوسط سطح تحصیلات پایین‌تر یا نرخ بیکاری بالاتری داشته باشند، مدل این آسیب‌پذیری را به درستی به متغیرهای «تحصیلات» و «اشتغال» نسبت می‌دهد، نه صرفاً به جنسیت. بنابراین، این نتیجه‌گیری نه تنها اهمیت جنسیت را نفی نمی‌کند، بلکه نشان می‌دهد که نابرابری‌های جنسیتی در فقر، خود را از کanal‌های ساختاری بازتولید می‌کنند و سیاست‌گذاری مؤثر باید همین کanal‌ها را هدف قرار دهد.

## نکته‌ی دوم | عدم معناداری آماری متغیر وضعیت تأهل سرپرست خانوار

یکی از نتایج قابل توجه، عدم معناداری آماری متغیر «وضعیت تأهل سرپرست خانوار» در تمامی دسته‌ها (مجرد، مطلقه و بیوه) در مقایسه با گروه مرجع (متأهل) است. این یافته به معنای بی‌اهمیت بودن وضعیت زناشویی در زندگی واقعی نیست، بلکه نشان می‌دهد که تأثیر اقتصادی آن به طور مؤثری توسط سایر متغیرهای کلیدی موجود در مدل جذب و نمایندگی شده است. به عنوان مثال، آسیب‌پذیری اقتصادی که غالباً با بیوه بودن همراه است، به شدت با سن سرپرست خانوار همبستگی دارد؛ مدل حاضر با کنترل متغیر سن (به ویژه گروه بالای ۷۰ سال)، این اثر را به درستی به عامل سن نسبت می‌دهد، نه صرفاً به وضعیت تأهل. به طور مشابه، فشار اقتصادی بر یک سرپرست خانوار مجرد یا مطلقه که تنها نان‌آور است، بسیار دقیق‌تر توسط متغیر «نسبت واستگی اشتغال» اندازه‌گیری می‌شود تا با برچسب وضعیت زناشویی. در نتیجه، مدل با موفقیت از شناسایی‌های سطحی فراتر رفته و نشان می‌دهد که این سازوکارهای اقتصادی زیربنایی (مانند سن، ساختار خانوار و وضعیت اشتغال) هستند که عوامل تعیین‌کننده‌تری نسبت به خود برچسب وضعیت تأهل محسوب می‌شوند.

## نکته‌ی سوم | نقش اندازه‌ی خانوار در حضور متغیر نسبت واستگی

اگرچه در ادبیات اقتصادی، «اندازه‌ی خانوار» همواره به عنوان یکی از عوامل کلیدی تعیین‌کننده فقر شناخته می‌شود، ضرایب مربوط به دسته‌های مختلف این متغیر در مدل حاضر به سطح معناداری آماری نرسیده‌اند. این یافته‌ی به ظاهر متناقض، در واقع نشان‌دهنده‌ی یک تفکیک تحلیلی دقیق است. دلیل اصلی این پدیده، حضور متغیر بسیار مهم و تحلیلی «نسبت واستگی اشتغال» در مدل است که خود به طور مستقیم با ابعاد و ساختار خانوار مرتبط است. در حقیقت، مدل نشان می‌دهد که صرف تعداد اعضای خانوار به تنها یکی عامل تعیین‌کننده ریسک فقر نیست، بلکه ساختار اقتصادی آن خانوار (یعنی نسبت اعضای بدون درآمد به اعضای دارای درآمد) است که اهمیت دارد. متغیر «نسبت واستگی» این مفهوم را بسیار دقیق‌تر از متغیر خام «اندازه‌ی خانوار» اندازه‌گیری می‌کند. بنابراین، مدل به درستی اثر فشار اقتصادی ناشی از بزرگ شدن خانوار را به این متغیر دقیق‌تر نسبت داده و پس از کنترل آن، متغیر ساده‌تر اندازه‌ی خانوار دیگر قدرت توضیحی مازادی نداشته است. این یافته تأکید می‌کند که سیاست‌های حمایتی نباید صرفاً بر اساس تعداد اعضا، بلکه باید با در نظر گرفتن نسبت نان‌آوران به کل اعضای خانوار طراحی شوند.

ارزیابی مدل | آزمون هاسمر-لمشو یک نتیجه آماری معنادار ( $p-value < 0.05$ ) را نشان می‌دهد که به وجود تفاوت میان احتمالات پیش‌بینی شده و نرخ‌های واقعی فقر در زیرگروه‌ها اشاره دارد. با وجود آنکه بخشی از این نتیجه می‌تواند به حساسیت بالای این آزمون در نمونه‌های آماری بزرگ ( $N=5823$ ) نسبت داده شود، اما این عدم کالیبراسیون دقیق، خود می‌تواند شاهدی بر محدودیت‌های ذاتی مدل لجستیک باشد. به طور مشخص، این مدل یک رابطه خطی بین متغیرها و لگاریتم شانس فقر را مفروض می‌گیرد، در حالی که بسیاری از روابط در دنیای واقعی (مانند اثر سن) غیرخطی هستند. این عدم انعطاف‌پذیری مدل در ثبت الگوهای پیچیده و غیرخطی، محتمل‌ترین دلیل برای ناتوانی آن در کالیبراسیون دقیق پیش‌بینی‌هاست. اینجاست که ضرورت حرکت به سوی مدل‌های انعطاف‌پذیرتر مانند GAM آشکار می‌شود. بنابراین، ضمن تأیید قدرت تفکیک بالای مدل رگرسیون لجستیک ( $R^2 = 0.28$ )، نتیجه‌ی آزمون هاسمر-لمشو به عنوان یک چراغ راهنمای عمل کرده و ما را به سمت استفاده از GAM برای کشف دینامیک‌های غیرخطی فقر هدایت می‌کند.

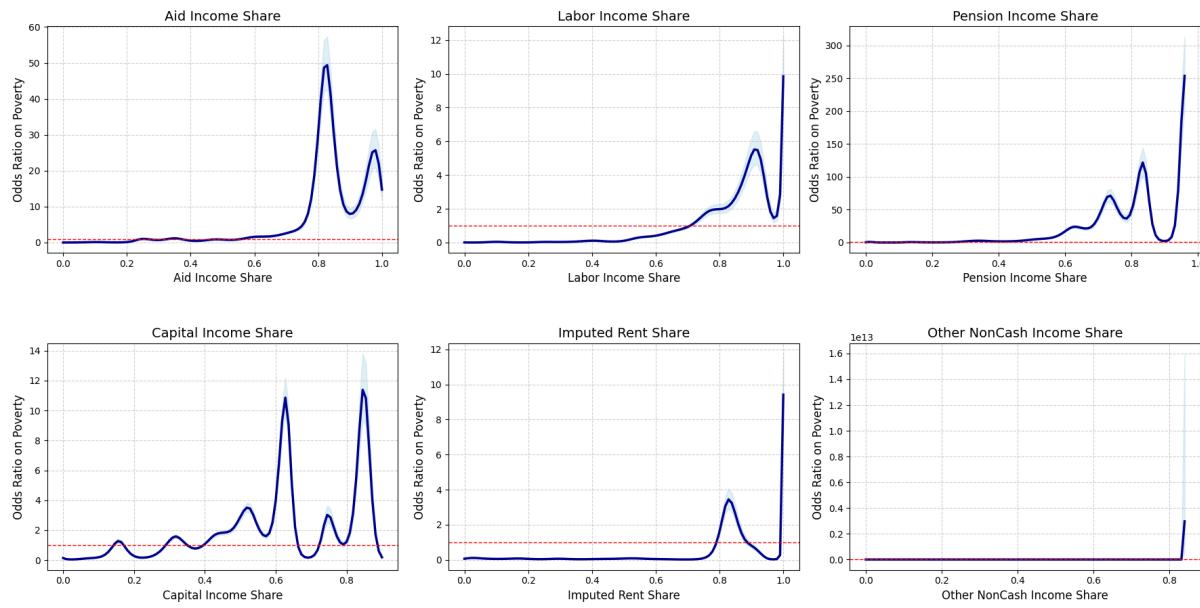
این محدودیت، ما را به این نتیجه رساند که انتکای صرف به ضرایب این مدل می‌تواند به تفاسیر ساده‌انگارانه منجر شود. از این رو، برای درک عمیق‌تر و دقیق‌تر دینامیک‌های فقر، تحلیل خود را با استفاده از یک مدل انعطاف‌پذیرتر، یعنی مدل جمعی تعمیم‌یافته (GAM)، ادامه دادیم تا روابط غیرخطی موجود در داده‌ها را به درستی مدل‌سازی و تفسیر کنیم.

## ۲.۵. مدل جمعی تعمیم‌یافته

برای غلبه بر محدودیت رگرسیون لاجستیک، از مدل جمعی تعمیم‌یافته (GAM) استفاده شد که فرض خطی بودن را کنار گذاشت و به داده‌ها اجازه می‌دهد تا خود، شکل رابطه بین هر متغیر و احتمال فقر را تعیین کنند. خروجی اصلی و بصری این مدل، نمودارهای وابستگی جزئی<sup>۱۳</sup> (نمودارهای ۱ و ۲) هستند که برای هر متغیر به صورت جداگانه ترسیم می‌شوند. این منحصراً به ما اجازه می‌دهد تا بینیم چگونه با تغییر یک متغیر، تأثیر آن بر شانس فقر به صورت غیرخطی تغییر می‌کند، در حالی که اثر سایر متغیرها در مدل ثابت نگه داشته شده است. مدل GAM با داشتن  $R^2 = 0.41$  Pseudo- $R^2$  قدرت تبیین‌کنندگی پسیار خوبی دارد و با توجه به کنترل درجه آزادی مؤثر ( $EDoF \approx 148$ ) از طریق جریمه‌سازی، بدون علائم جدی بیش‌برازش حاصل شده است.

### بحث اول: سهم درآمدهای حمایتی (کمک‌های نقدی و بازنشستگی)

نماودارهای سهم درآمد از کمک‌های نقدی (Aid Income Share) و سهم درآمد از حقوق بازنشستگی (Pension Income Share) الگوی بسیار مشابه و معناداری را نشان می‌دهند. تا زمانی که سهم این درآمدها در سبد درآمدی خانوار پایین تا متوسط است (تقریباً زیر ۶۰٪)، تأثیر معناداری بر افزایش شانس فقر ندارند و نسبت شانس نزدیک به ۱ است. اما با افزایش این سهم و نزدیک شدن به وابستگی کامل (بالای ۸۰٪)، شانس قرار گرفتن در زیر خط فقر به صورت نمایی و شدید افزایش می‌یابد. در نتیجه، فقر به صرف دریافت کمک یا بازنشستگی نیست، بلکه به دلیل وابستگی شدید و تقریباً انحصاری به این منابع است. این مدل یک «نقشه آستانه» را شناسایی می‌کند که پس از آن، این منابع درآمدی که ذاتاً حمایتی هستند، به شاخصی برای آسیب‌پذیری شدید اقتصادی تبدیل می‌شوند. این یک رابطه کاملاً غیرخطی است که مدل لجستیک قادر به ثبت آن نبود.



نمودار ۱ – تأثیر سهم انواع درآمدها بر فقر در مدل GAM

## بحث دوم: سهم درآمد از کار و اجاره ضمنی

در نمودار سهم درآمد از کار (Labor Income Share)، ابتدا با افزایش سهم درآمد از کار، شانس فقر کاهش می‌یابد (نسبت شانس کمتر از ۱) که منطقی است و نقش محافظتی اشتغال را نشان می‌دهد. اما به طرز شگفت‌انگیزی، وقتی این سهم از حدود ۸۰٪ فراتر می‌رود، شانس فقر به شدت افزایش می‌یابد. این الگو به احتمال زیاد در حال شناسایی گروه «فقرای شاغل» است. یعنی خانوارهایی که تقریباً تمام درآمدشان از محل کار است، اما به دلیل دستمزد پایین یا فقدان سایر منابع درآمدی (مانند درآمد سرمایه‌یا نداشتن ملک)، همچنان زیر خط فقر باقی می‌مانند. مدل به درستی نشان می‌دهد که صرفاً شاغل بودن کافی نیست، بلکه میزان درآمد و تنوع آن اهمیت دارد.

نمودار سهم درآمد از اجاره ضمنی (Imputed Rent Share) نیز الگویی بسیار گویا دارد. برای بخش بزرگی از دامنه، این متغیر اثر محافظتی اندکی دارد. اما زمانی که سهم آن به نزدیکی ۱۰۰٪ می‌رسد، شانس فقر به شدت و به صورت عمودی افزایش می‌یابد. این نمودار به طور دقیق گروه «ثروتمند از نظر دارایی، اما فقیر از نظر درآمدی» را شناسایی می‌کند که در بحث قبلی به آن اشاره شد. این‌ها خانوارهایی هستند که مالک مسکن خود هستند (و ارزش اجاره ضمنی بالایی دارند) اما تقریباً هیچ جریان درآمد نقدی دیگری ندارند.

## بحث سوم: سهم درآمدهای سرمایه‌ای و غیرنقدی

نمودار سهم درآمد از سرمایه (Capital Income Share) الگوی بسیار پیچیده و پرنوسانی دارد. اگرچه در ابتدا اثر محافظتی دارد (نسبت شانس کمتر از ۱)، اما در نقاط مختلف جهش‌های بزرگی را نشان می‌دهد که تفسیر آن‌ها دشوار است. این الگو به احتمال زیاد تحت تأثیر تعداد کم مشاهدات در سطوح بالای سهم درآمد از سرمایه قرار گرفته است. بازه‌های اطمینان (ناحیه آبی کمرنگ) در نقاط اوج، بسیار عریض‌تر هستند که نشان‌دهنده عدم قطعیت بالای مدل در این نواحی است. ارائه یک تفسیر اقتصادی یکپارچه برای این نوسانات دشوار است.

نمودار سهم سایر درآمدهای غیرنقدی (Other NonCash Income Share) برای تقریباً کل دامنه، اثری نامشخص، اما در انتهای طیف (بالای ۸۰٪) یک جهش بسیار شدید و غیرعادی نشان می‌دهد.

❖ افت و خیزهای شدید در این نمودار و سایر نمودارها (به خصوص در انتهای) ممکن است به دلیل تعداد کم داده در آن نقطه باشد و باید با احتیاط تفسیر شود.

#### بحث چهارم: عوامل جمعیت‌شناسخی (سن سرپرست و نسبت وابستگی)

نمودار سن سرپرست (Head Age) یک الگوی نزولی در چرخه عمر فقر را نشان می‌دهد. ریسک فقر در سنین جوانی (حدود ۲۰ سالگی) به شدت بالاست، به طوری که شانس فقر برای یک سرپرست ۲۰ ساله بیش از ۳۰ برابر یک فرد میانسال است. این امر به وضوح آسیب‌پذیری جوانان در ابتدای مسیر شغلی، با درآمدهای پایین و عدم ثبات شغلی را نشان می‌دهد. با افزایش سن، این ریسک به سرعت کاهش یافته و در بازه سنی ۳۰ تا ۶۵ سالگی به کمترین میزان خود می‌رسد و نسبت شانس در نزدیکی یا زیر ۱ باقی می‌ماند. در سنین سالمندی (بالای ۷۰ سال)، ریسک فقر مجدداً کمی افزایش می‌یابد اما هرگز به شدت دوران جوانی نمی‌رسد. این الگو نشان می‌دهد که سیاست‌های حمایتی باید به طور ویژه بر جوانان سرپرست خانوار متوجه شوند.

نمودار نسبت وابستگی اشتغال (Dependency Ratio Employment) تأثیر مستقیم و نمایی ساختار خانوار بر فقر را به تصویر می‌کشد. تا زمانی که نسبت افراد غیرشاغل به شاغل در خانوار پایین است (کمتر از ۴)، افزایش آن تأثیر چندانی بر شانس فقر ندارد. اما پس از عبور از این آستانه، منحنی صعودی می‌شود. در نقطه اوج، زمانی که این نسبت به ۸ می‌رسد (یعنی به ازای هر فرد شاغل، ۸ نفر بدون درآمد وجود دارند)، شانس فقر به طور شدیدی به بیش از ۱۲۰ برابر افزایش می‌یابد. این یافته به طور قاطع نشان می‌دهد که صرفاً تعداد اعضای خانوار مهم نیست، بلکه ساختار اقتصادی آن و بار تکفل بر دوش افراد شاغل، عامل بحرانی در تعیین ریسک فقر است.

#### بحث پنجم: سرمایه انسانی (تحصیلات و اشتغال سرپرست)

نمودار سطح تحصیلات سرپرست (Head Education Level) یک رابطه شبیدار و تقریباً خطی در مقیاسنگاریتمی را نشان می‌دهد و به قوی‌ترین شکل ممکن، نقش محوری تحصیلات را به عنوان یک سپر محافظ در برابر فقر تأیید می‌کند. شانس فقر برای یک سرپرست خانوار بی‌سواد، ۷ برابر بیشتر از یک فرد دارای دیپلم (گروه مرجع با نسبت شانس ۱) است. با هر پله افزایش در سطح تحصیلات، این شانس به شدت کاهش می‌یابد: شانس فقر برای افراد زیر دیپلم حدود ۳/۲ برابر، برای دارندگان مدرک کارشناسی ۳/۰ برابر (یعنی ۷۰٪ کمتر) و برای دارندگان تحصیلات تکمیلی تنها ۱/۰ برابر (یعنی ۹۰٪ کمتر) از گروه مرجع است. این شبیده نشان می‌دهد که بازده اقتصادی تحصیلات در کاهش فقر، بسیار چشمگیر و تعیین‌کننده است.

نمودار وضعیت فعالیت سرپرست (Head Activity Status) نیز به وضوح اهمیت نوع اشتغال را نشان می‌دهد. گروه «سایر» که شامل کارگران فصلی و مشاغل غیررسمی است، با نسبت شانس ۱/۷۵، آسیب‌پذیرترین گروه محسوب می‌شود. بیکاران (Unemployed) و خانه‌داران (Housekeeper) نیز به ترتیب با نسبت شانس ۱/۴ و ۱/۱، شانس بیشتری برای فقر نسبت به شاغلان (Employed) دارند. نکته‌ی جالب، گروه «دارای درآمد بدون کار» (Income\_without\_Work) است که با نسبت شانس حدود ۰/۵۵، حتی از شاغلان نیز وضعیت بهتری دارند. این نشان می‌دهد که داشتن جریان درآمدی پایدار غیرشغلی (مانند مستمری یا سود سرمایه) یک عامل

محافظتی بسیار قوی در برابر فقر است. شاید با ترکیب مناسبی از درآمد شغلی و درآمد سرمایه یا بازنشستگی بتوان از خط فقر فاصله‌ی زیادی گرفت.

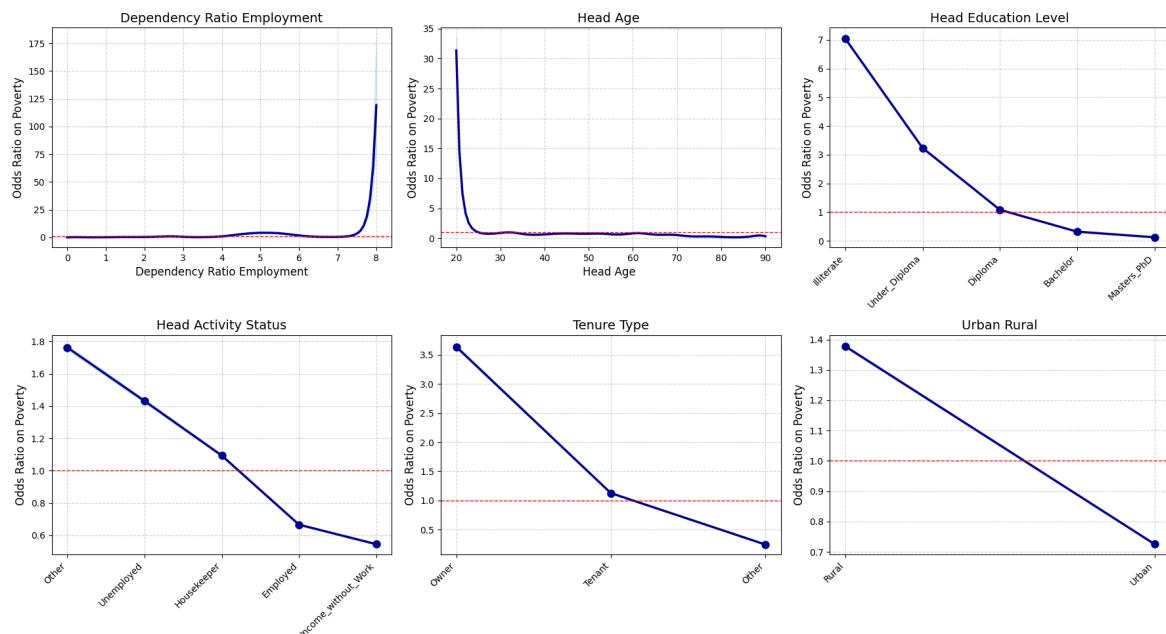
### بحث ششم: عوامل ساختاری (منطقه‌ی سکونت و نحوه‌ی تصرف مسکن)

نمودار نوع تصرف مسکن یکی از پیچیده‌ترین و در نگاه اول، متناقض‌ترین یافته‌های مدل را به نمایش می‌گذارد. برخلاف تصور رایج، گروه «مالک (Owner)» با نسبت شانس حدود ۳/۶، بالاترین ریسک فقر را در میان دسته‌ها نشان می‌دهد. این در حالی است که گروه «مستأجر (Tenant)» نسبت شانس تقریباً ۱/۱، بسیار نزدیک به خط مبدا (نسبت شانس = ۱) قرار دارد و گروه «سایر» (شامل سکونت رایگان و...) با نسبت شانس تنها ۰/۲۵ (یعنی ۷۵٪ کمتر از مستأجران)، از کمترین ریسک برخوردار است.

این یافته به معنای آن نیست که مالکیت به خودی خود عامل فقر است، بلکه نشان‌دهنده‌ی توانایی مدل GAM در تفکیک اثرات پیچیده است. این نمودار را باید همزمان با نمودار «سهم درآمد از اجاره ضمنی» تفسیر کرد. مدل به درستی گروهی خاص از مالکان «ثروتمند از نظر دارایی، اما فقیر از نظر درآمدی» را شناسایی.

بنابراین، این نمودار به تنهایی نباید تفسیر شود. این الگو نشان می‌دهد که مدل، مالکان کم‌درآمد را به عنوان گروهی پر ریسک شناسایی می‌کند، در حالی که اثر حمایتی دارایی آن‌ها توسط متغیر دیگری به طور دقیق‌تری در مدل لحاظ شده است. این تفکیک تحلیلی، برتری مدل GAM را در شناسایی روابط چندوجهی و جلوگیری از نتایج سطحی به خوبی نشان می‌دهد.

نمودار سکونتگاه شهری/ روستایی (Urban / Rural) شکاف عمیق منطقه‌ای را به سادگی به تصویر می‌کشد. شانس قرار گرفتن در زیر خط فقر برای خانوارهای روستایی (Rural) حدود ۱/۳۵ برابر خانوارهای شهری (Urban) است. این یافته، حتی پس از کنترل تمام عوامل دیگر مانند تحصیلات، شغل، سن و ساختار درآمدی، همچنان معنادار است و نشان می‌دهد که «صرف سکونت در روستا» به خودی خود یک عامل ریسک مستقل برای فقر محسوب می‌شود که احتمالاً به دلیل دسترسی کمتر به زیرساخت‌ها و فرصت‌های اقتصادی است.



نمودار ۲ – تأثیر سایر متغیرها بر فقر در مدل GAM

## ۶. محدودیت‌ها و مسیر آینده

یکی از دستاوردهای کلیدی این پژوهش، جنبه روش‌شناسخی آن است. در حالی که مدل رگرسیون لاجستیک توانست عوامل اصلی مرتبط با فقر را شناسایی کند، آزمون نیکویی برازش، ناکارآمدی آن را در مدل‌سازی روابط پیچیده آشکار ساخت. در مقابل، مدل GAM با ارائه‌ی نگاهی باوضوح بالاتر، ماهیت واقعی و عمدتاً غیرخطی تأثیر این عوامل را به تصویر کشید. این مطالعه تأیید می‌کند که در پدیده‌های پیچیده اجتماعی مانند فقر، اتفاک به مدل‌هایی با فروض محدودکننده می‌تواند به برداشت‌های ساده‌انگارانه منجر شود و استفاده از مدل‌های نیمه‌پارامتریک و انعطاف‌پذیر برای کشف دینامیک‌های پنهان اقتصادی و اجتماعی ضروری است.

با این وجود، پژوهش حاضر با محدودیت‌هایی نیز مواجه است که باید در تفسیر نتایج مد نظر قرار گیرند. نخست، ماهیت مقطوعی داده‌ها موجب می‌شود که تحلیل‌ها صرفاً همبستگی‌ها را نشان دهند و امکان استنتاج رابطه‌ی علی میان متغیرها و فقر وجود نداشته باشد. برای مثال، نمی‌توان با قطعیت گفت که تحصیلات پایین‌علت فقر است یا هر دو، معلولِ شرایط زمینه‌ای دیگری هستند. دوم، تورش متغیرهای حذف شده<sup>۱۴</sup> یک نگرانی بالقوه است؛ متغیرهای مشاهده‌نشده‌ای مانند وضعیت سلامت اعضای خانوار، سطح بدھی، دسترسی به شبکه‌های حمایت غیررسمی و یا حتی شاخص‌های روان‌شناسخی مانند انگیزه و امید به آینده، می‌توانند بر وضعیت فقر تأثیرگذار باشند که در داده‌های فعلی در دسترس نبودند.

علاوه بر این، همان‌طور که در تحلیل نتایج مدل GAM مشاهده شد، الگوهای تأثیر برخی متغیرها بسیار پیچیده و پرونوسان بود. این موضوع نشان می‌دهد که حتی یک مدل انعطاف‌پذیر مانند GAM نیز ممکن است برای ثبت کامل تمام دینامیک‌های زیربنایی، نیازمند داده‌های بیشتر یا متغیرهای تعاملی دقیق‌تری باشد. این پیچیدگی‌ها، خود به عنوان یک یافته، بر ماهیت چندوجهی فقر تأکید کرده و نشان می‌دهند که جای کار بیشتری برای درک این روابط وجود دارد.

بر اساس این یافته‌ها و محدودیت‌ها، چندین مسیر برای تحقیقات آتی پیشنهاد می‌شود. نخست، استفاده از داده‌های پانل<sup>۱۵</sup> در صورت دسترسی، می‌تواند به تحلیل‌های علی و بررسی پویایی‌های فقر در طول زمان کمک شایانی کند. دوم، همان‌طور که اشاره شد، تحلیل جداگانه بر روی زیرگروه‌های خاص مانند خانوارهای زن‌سرپرست یا سالمندان، می‌تواند سازوکارهای منحصر به فرد فقر در این گروه‌ها را آشکار سازد؛ بهویژه آنکه متغیرهایی مانند جنسیت و وضعیت تأهل در مدل کلی معنادار نبودند. در نهایت، پیشنهاد می‌شود در مطالعات آینده، اثرات تعاملی میان متغیرهای کلیدی (مانند تعامل سطح تحصیلات با محل سکونت) با استفاده از مدل‌های پیشرفته‌تری مانند جنگل‌های تصادفی یا GAM‌های تعاملی بررسی شود تا درک دقیق‌تری از چگونگی تأثیر همزمان عوامل مختلف بر ریسک فقر به دست آید.

## ۷. نتیجه‌گیری و فرصت‌های آینده

این پژوهش با بهره‌گیری از یک مدل جمعی تعمیم‌یافته (GAM)، تصویری چندبعدی و دقیق از عوامل تعیین‌کننده فقر نسبی در خانوارهای تهرانی ارائه داد. نتایج نشان می‌دهد که سرمایه انسانی، به ویژه سطح تحصیلات سرپرست خانوار، قدرتمندترین سپر محافظ در برابر فقر است. این مدل با آشکار ساختن روابط غیرخطی، نشان داد که آسیب‌پذیری اقتصادی صرفاً یک پدیده خطی نیست؛ برای مثال، ریسک فقر در سنین جوانی به شدت بالاست و با افزایش سن کاهش می‌یابد (الگوی چرخه عمر)، و وابستگی شدید به درآمدهای حمایتی تها پس از عبور از یک نقطه آستانه مشخص، به شاخصی برای فقر شدید تبدیل می‌شود. این یافته‌ها بر ضرورت طراحی سیاست‌های اجتماعی هدفمند که به جای تمرکز بر برچسب‌های کلی، بر مراحل خاص چرخه زندگی (مانند جوانان) و ساختارهای اقتصادی خانوار (مانند نسبت وابستگی اشتغال) متمرکز شوند، تأکید می‌کند.

علاوه بر این، مدل GAM با موفقیت توانست پارادوکس‌های ظاهری در داده‌ها را تبیین کند. شناسایی گروه‌های خاصی مانند «فقرای شاغل» (که علی‌رغم اتکای کامل به درآمد کاری، فقیر هستند) و «مالکان فقیر از نظر درآمدی» (که دارای ثروت غیرنقدی اما فاقد جریان درآمدی کافی هستند)، نشان‌دهنده توانایی مدل در فراتر رفتن از تحلیل‌های سطحی است. این مطالعه تأیید می‌کند که فقر پدیده‌ای تکبعدي نیست و عواملی مانند تنوع منابع درآمدی و نسبت دارایی به درآمد، نقشی حیاتی در تعیین وضعیت اقتصادی خانوار ایفا می‌کنند. در نهایت، عدم معناداری متغیرهایی مانند جنسیت یا وضعیت تأهل، نه به معنای بی‌اهمیت آن‌ها، بلکه نشان‌دهنده جذب اثراشان توسط متغیرهای اقتصادی ساختاری‌تر (مانند اشتغال و تحصیلات) در مدل است. این یافته خود، اهمیت پرداختن به نابرابری‌های ریشه‌ای در دسترسی به فرصت‌ها را به عنوان مؤثرترین راهبرد برای مبارزه با فقر برجسته می‌سازد.

## ۸. پیوست

در پیش‌پردازش داده‌ها، ابتدا یک برنامه‌ی پایتون نوشته شد که همه‌ی جداول سال‌های ۱۳۹۹ تا ۱۴۰۱ را از کتابخانه‌ی HBSIR دریافت کند. سپس مطابق موضوع پژوهه، متغیرهای مورد نیاز را شناسایی کردیم و با یک برنامه‌ی پایتون دیگر، آن‌ها را از جداول مربوطه استخراج کردیم. فرآیند انتخاب متغیرها با استناد به مقالات علمی انجام شد.

پس از تجمعی روش‌های مقالات و گرفتن بازخورد از مدل رگرسیون لاجستیک تصمیم گرفتیم دسته‌بندی برخی متغیرها را تغییر دهیم تا ضمن حفظ دقت، مدل پایدار شود. دسته‌بندی متغیرهای چندکلاسه به شرح زیر است:

- تحصیلات: بی‌سواد، زیر دیپلم، دیپلم، لیسانس، فوق لیسانس
- اشتغال: دارای درآمد بدون کار، شاغل، خانه‌دار، بی‌کار، سایر
- نحوه‌ی تصرف مسکن: مالک، مستأجر، سایر
- تعداد اعضای خانوار: ۱، ۲، ۳، ۴، ۵، ۶، ۷، بالای ۷

هم‌چنین از روی جدول ریز درآمدها، شش متغیر جدید معادل با سهم آن‌ها از درآمد کل ساختیم. جزئیات و توضیحات بیشتر آن در جدول ۲ موجود است.

جدول ۲ – تقسیم‌بندی انواع درآمدهای خانوار

گروه درآمدی اصلی (متغیر مدل)	زیرشاخه‌های درآمدی خام (متغیرهای اولیه)	شرح گروه
Labor Income (درآمد از کار)	Cash_Private, Cash_Public Cash_Home_Production Cash_NonAgricultural Cash_Agricultural Cash_Cooperative	شامل کلیه درآمدهای نقدی حاصل از اشتغال در بخش‌های خصوصی، دولتی، تعاونی و همچنین فعالیت‌های مستقل کشاورزی و غیرکشاورزی
Pension Income (درآمد بازنیستگی)	Cash_Retirement	شامل کلیه درآمدهای نقدی حاصل از مستمری و حقوق بازنیستگی
Capital Income (درآمد از سرمایه)	Cash_Rent Cash_Interest	شامل کلیه درآمدهای نقدی ناشی از دارایی‌ها، مانند اجاره‌بهای دریافتی و سود سپرده‌های بانکی
Aid Income (درآمدهای حمایتی)	Cash_Aid Cash_Subsidy Cash_Transfer NonCash_Public NonCash_Donation	شامل کلیه یارانه‌های نقدی، کمک‌های نقدی و غیرنقدی دریافتی از نهادهای دولتی و خصوصی، و انتقالات بین خانوارها
Imputed Rent (اجاره ضمنی)	NonCash_ImputedRent_Ownership NonCash_ImputedRent_Mortgage	ارزش غیرنقدی استفاده از مسکن ملکی (مالکیت قطعی یا در رهن) که به عنوان یک درآمد غیرنقدی در نظر گرفته شده است
Other NonCash Income (سایر درآمدهای غیرنقدی)	NonCash_Private NonCash_Cooperative NonCash_Agricultural NonCash_NonAgricultural NonCash_HomeProduction	شامل سایر درآمدهای غیرنقدی مانند ارزش محصولات کشاورزی یا تولیدات خانگی برای مصرف شخصی و کمک‌های غیرنقدی از بخش خصوصی