

آموزش

تحلیل مسیر

با نرم افزار SPSS

(Path Analysis)

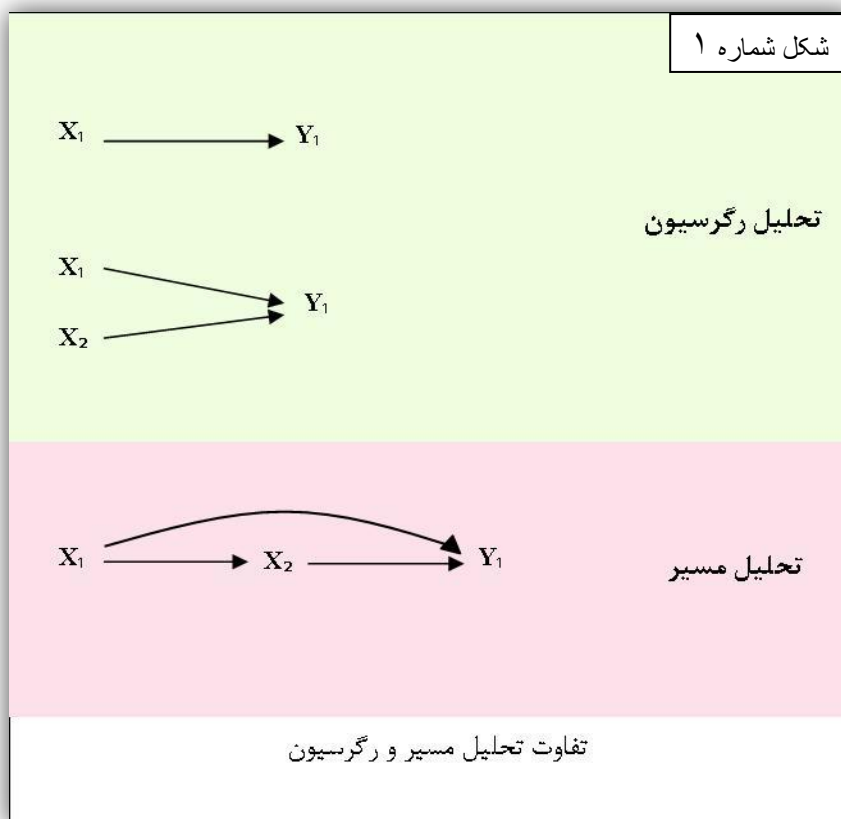
تدوین: رامین کریمی

www.Kharazmi-Statistics.ir

تحلیل مسیر (Path Analysis) نخستین بار توسط یک بیولوژیست به نام سوول رایت (Swell Wright) در سال ۱۹۳۴ در ارتباط با تجزیه کردن مجموع همبستگی بین دو متغیر در یک سیستم علی معرفی گردید. این روش تعمیم یافته روش رگرسیون چند متغیره در ارتباط با تدوین مدل های علی است. تکنیک تحلیل مسیر بر اساس فرض ارتباط بین متغیرهای مستقل و وابسته استوار است. این روش بر استفاده ابتکاری از نمودار تصویری که به دیاگرام مسیر (Path Diagram) است، تاکید خاص دارد.

در روش تحلیل مسیر، علاوه بر تاثیر مستقیم، امکان شناسایی تاثیرات غیرمستقیم هر یک از متغیرهای مستقل بر متغیر وابسته نیز وجود دارد. به همین خاطر، در تحلیل مسیر، با چندین معادله خط رگرسیونی استاندارد شده مواجه هستیم، در حالی که در تحلیل رگرسیون، تنها با یک معادله خط رگرسیونی استاندارد شده داریم.

در واقع تفاوت تحلیل مسیر و رگرسیون در توانایی تحلیل مسیر در سنجش تاثیرات غیرمستقیم (علاوه بر تاثیرات مستقیم) است. این تفاوت در شکل زیر نمایش داده شده است.



نکته: در تحلیل مسیر، فلش ها مشخص کننده تاثیرات علی از متغیرهای مستقل به سمت متغیرهای وابسته میانی و نهایی می باشند.

تحلیل مسیر برای تبیین دقیق تر روابط علی متغیرها، به تجزیه همبستگی بین متغیرها می پردازد. از طریق این تجزیه، اثرات مستقیم و غیر مستقیم یک متغیر، بر متغیر دیگر مشخص می گردد. آن چه که ضرایب همبستگی نشان می دهند، اثرات مستقیم و غیرمستقیم متغیرهاست. به عبارت دیگر:

$$\text{اثرات غیر مستقیم} + \text{اثرات مستقیم} = \text{همبستگی}$$

مفاهیم کلیدی در تحلیل مسیر

مفروضات

تحلیل مسیر بر مبنای مفروضاتی استوار است که عمده ترین آن ها به شرح ذیل اند:

۱- روابط بین متغیرهای موجود در مدل، خطی، جمع پذیر و علی است و روابط انحنایی و تعاملی ملحوظ نمی گردد.

۲- متغیرهای باقیمانده با همدیگر و با متغیرهایی که قبل از آن در مدل قرار گرفته اند، همبسته نیستند.

۳- جریان علیت در دستگاه یک طرفه می باشد و علیت متقابل بین متغیرها ملحوظ نمی شود.

نمودار یا دیاگرام مسیر

مهم ترین بخش تحلیل مسیر، طراحی و آزمون نمودار مسیر است که از چند جزء اساسی تشکیل شده است. نمودار مسیر، در واقع، یک مدل ساختاری پیشین یا پیش تجربی با مجموعه معادله ساختاری است که روابط علی ممکن بین متغیرها را توصیف می کند. این نمودار مسیر، همواره پس از مرور مبانی نظری و تدوین چارچوب نظری انتخابی تحقیق توسط محقق طراحی می شود که در نهایت در تحلیل مسیر مورد آزمون تجربی قرار می گیرد.

انواع متغیر در نمودار مسیر

۱) متغیرهای درونی: در تحلیل مسیر، متغیرهای درونی شامل متغیر وابسته و متغیر مستقل است.

الف) متغیر وابسته: در تحلیل مسیر با دو دسته کلی از متغیر وابسته سروکار داریم:

۱- متغیر وابسته نهایی: متغیری که در نهایت تمامی تحلیل های تحقیق مبنی بر تاثیرات متغیرهای مستقل باید روی آن انجام شود (در مثال ما، میزان تماشای تلویزیون یا TV).

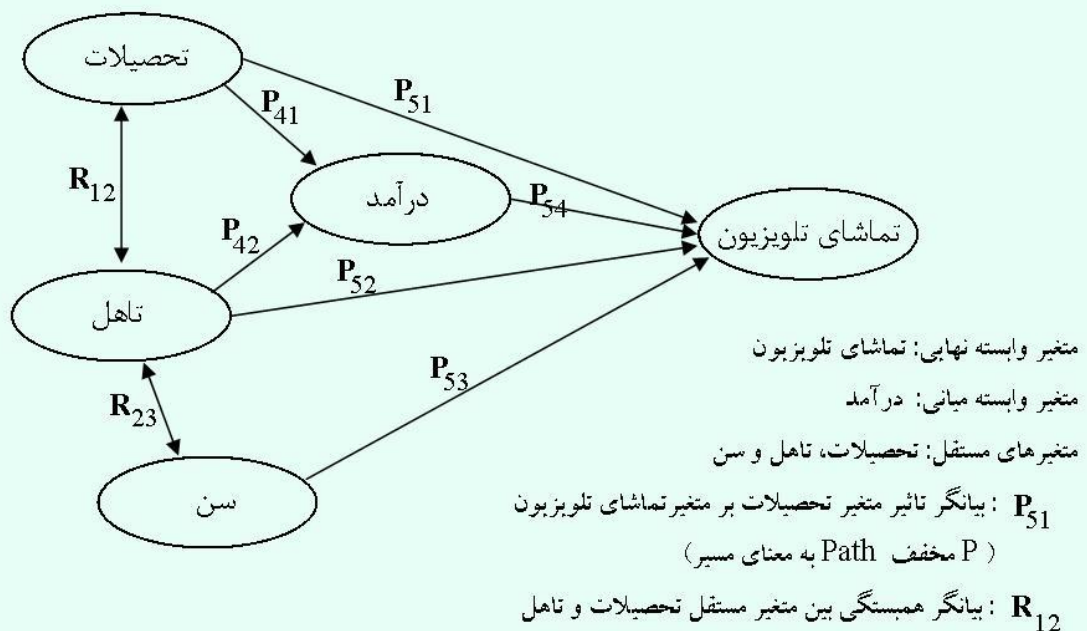
۲- متغیر وابسته میانی: در واقع همان متغیر مستقل است که در برخی مراحل تحلیل مسیر، نقش یک متغیر وابسته میانی یا واسطه را ایفا کرده و تاثیر سایر متغیرهای مستقل مدل بر روی آن آزمون می شود (در مثال ما متغیر درآمد یا Income).

ب) متغیر مستقل: متغیری که بر متغیرهای وابسته اعم از نهایی و میانی تاثیر می گذارد و به عبارتی از هیچ متغیری تاثیر نمی پذیرد. در مثال ما سایر متغیرها که شامل سن، وضعیت تاهل و تحصیلات می شوند متغیر مستقل هستند (شکل شماره ۲).

نکته: در هنگام آزمون مدل نظری، به بررسی میزان همبستگی دوجانبه بین متغیرهای می پردازیم تا پی ببریم آیا متغیرهای مستقل با همدیگر همبستگی دارند یا خیر. این رابطه از نوع همبستگی است و نمی توان آن را علی فرض کرد. وجود همبستگی بالا بین متغیرهای مستقل نشان می دهد که متغیرهایی که برای این پژوهش انتخاب شده اند دارای پشتوانه تئوریک هستند و افزایش یا کاهش در هر یک از این متغیرها، با افزایش یا کاهش در متغیر مقایسه شده با آن همراه می باشد.

نکته: متغیر وضعیت تاهل به صورت اسمی (مجرد = ۱ و متاهل = ۲) تعریف شده است. بنابراین پیش از اجرای رگرسیون، متغیرهای اسمی می بایست به طور تصنعی به فاصله ای تبدیل شوند. در نتیجه در SPSS متغیر دیگری تعریف کرده و جهت تبدیل متغیر اسمی وضعیت تاهل به متغیر فاصله ای، به کلیه افراد مجرد کد ۰ و به تمامی افراد متاهل کد ۱ می دهیم و این متغیر جدید را وارد تحلیل رگرسیون می کنیم.

شکل شماره ۲



ارزیابی مناسب بودن مدل: در تحلیل مسیر، برای ارزیابی مدل از آماره R^2 استفاده می شود. این آماره مقدار واریانس متغیر وابسته را نشان می دهد که متغیرهای مستقل توانسته اند آن را تبیین کنند. در واقع، R^2 نشان دهنده این است که مدل تا چه اندازه برازنده مجموعه ای از داده هاست. بنابراین هرچه مقدار R^2 بالاتر باشد، مدل قوی تر است و برعکس مقدار پایین R^2 دلالت بر ضعف مدل دارد که باید مدل دیگری بین متغیرها بسازیم که تبیین کننده واریانس بیش تری باشد.

آزمون ضرایب مسیر

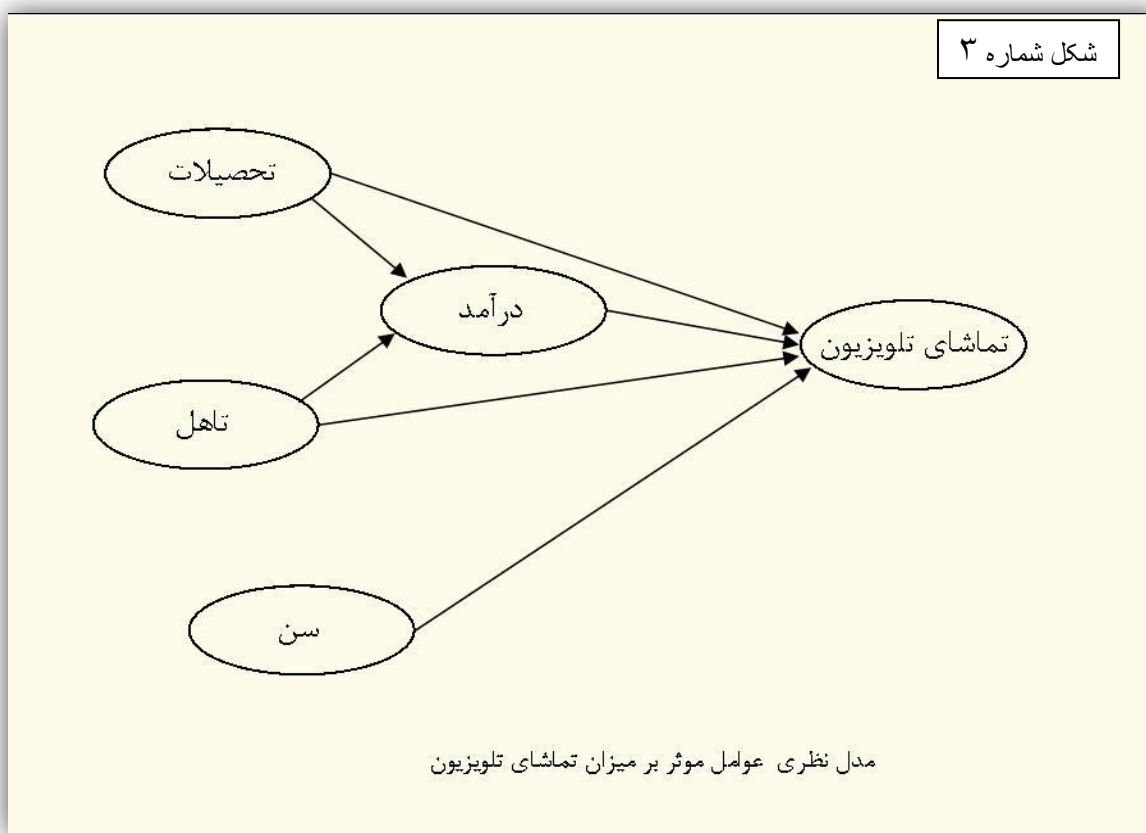
در تحلیل مسیر یک مدل نظری را به آزمون می گذاریم که در نهایت با اجرای تحلیل، این مدل نظری (که از چارچوب نظری به دست می آید) باید به یک مدل تجربی (که از تحلیل مسیر به دست می آید) منتهی شود. بنابراین طبیعی است که همواره ساخت روابط علی بین متغیرها در مدل تجربی با مدل نظری فرق دارد. اما چه متغیرهایی باید در مدل باقی بمانند و چه متغیرهایی باید از مدل حذف شوند؟

قاعده کلی این است که در اجرای تحلیل مسیر ، متغیرهایی که مقدار بتا یا ضریب مسیر (Beta) آن ها در سطح خطای کوچک تر از ۰.۵ / معنی دار نشود، از مدل حذف می کنیم. بنابراین، مسیر این متغیرها نیز از مدل حذف می شود. البته هر چه حجم نمونه بزرگ تر باشد، احتمال این که ضرایب مسیر معنی دار شوند، بیش تر است.

نحوه اجرا:

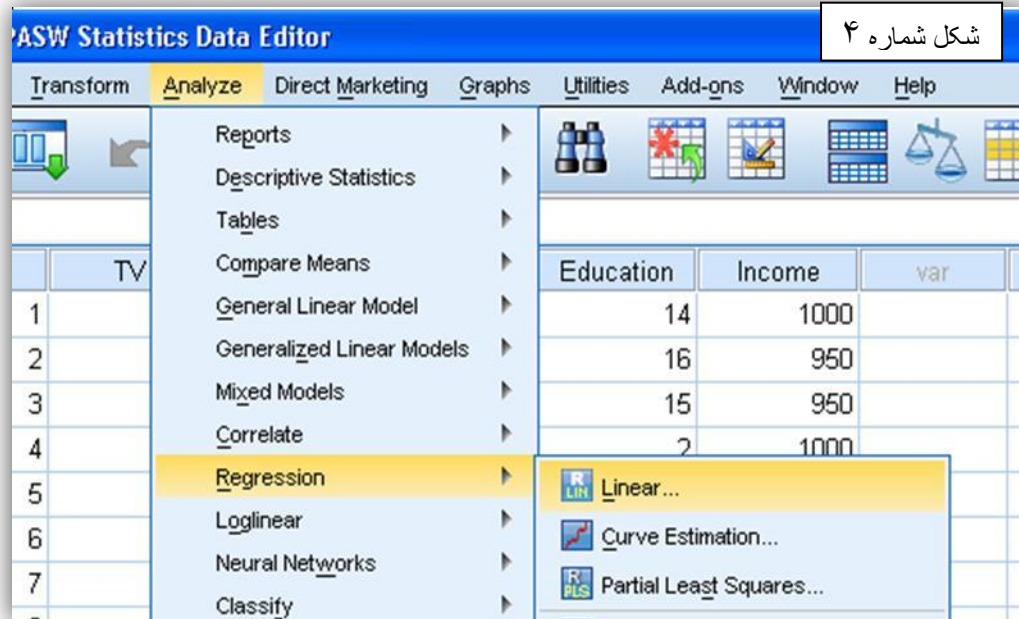
ما در این مثال پس از مطالعه مبانی نظری و نیز پیشینه تجربی مدل نظری زیر را تنظیم کرده ایم (البته این نمودار مسیر و داده ای مورد استفاده در آن ساختگی است و تنها برای آموزش مراحل و نحوه اجرای تحلیل مسیر به کار گرفته شده است).

این مدل عوامل موثر بر میزان تماشای تلویزیون (TV) را نشان می دهد. عوامل مورد نظر ما عبارت اند از: تحصیلات (Education)، وضعیت تاهل (Married)، سن (Age) و متغیر درآمد (Income). مدل نظری و مفهومی مد نظر ما به صورت زیر است.



مراحل زیر را جهت آزمون مدل نظری اجرا می کنیم.

۱- دستور **Analyze > Regression > Linear** را اجرا می کنیم.



۲- متغیر وابسته (TV) را وارد کادر **Dependent** و تمامی متغیرهای مستقل را وارد کادر **Independent** می کنیم. بر روی **OK** کلیک می کنیم. با این کار رگرسیون میزان بتای مسیرهای P_{51} و P_{52} و P_{53} و P_{54} (شکل شماره ۲) را به ما می دهد، اما هنوز میزان بتای سایر مسیرها به دست نیامده است.

Linear Regression شکل شماره ۵

Case

Married

Age

Education

Income

Dependent: TV

Block 1 of 1

Independent(s): Married, Age, Education

Method: Enter

Selection Variable:

Case Labels:

WLS Weight:

Statistics...

Plots...

Save...

Options...

Bootstrap...

OK Paste Reset Cancel Help

Coefficients^a جدول شماره ۱

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
1 (Constant)	4.597	.498		9.223	.000
Married	-.365	.079	-.153	-4.636	.000
Age	.004	.006	.019	.625	.533
Education	.001	.010	.003	.092	.927
Income	-.004	.000	-.868	-24.807	.000

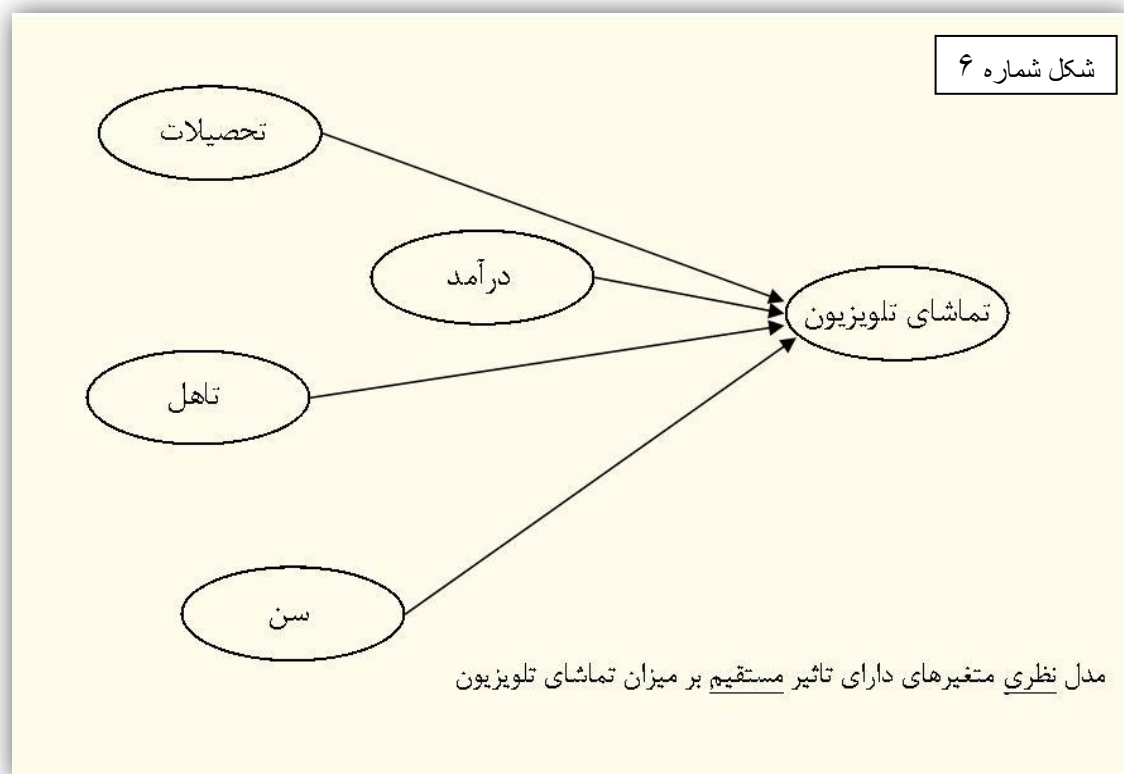
a. Dependent Variable: TV

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.962 ^a	.925	.923	.3294

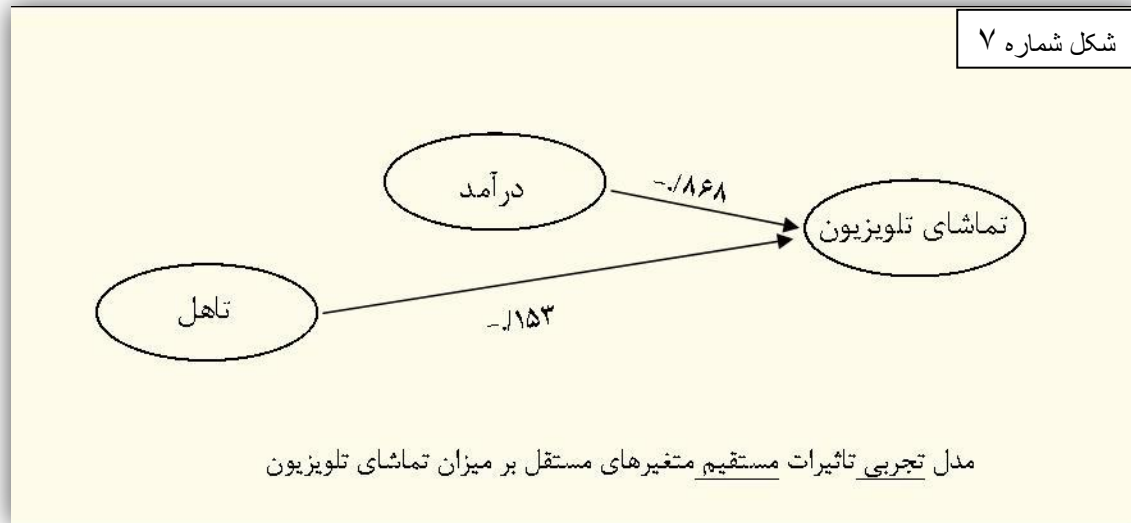
a. Predictors: (Constant), Income, Age, Married, Education

جدول شماره ۲

مدل نظری اثرات مستقیم متغیرهای مستقل بر متغیر وابسته نهایی در مرحله اول به صورت زیر است:



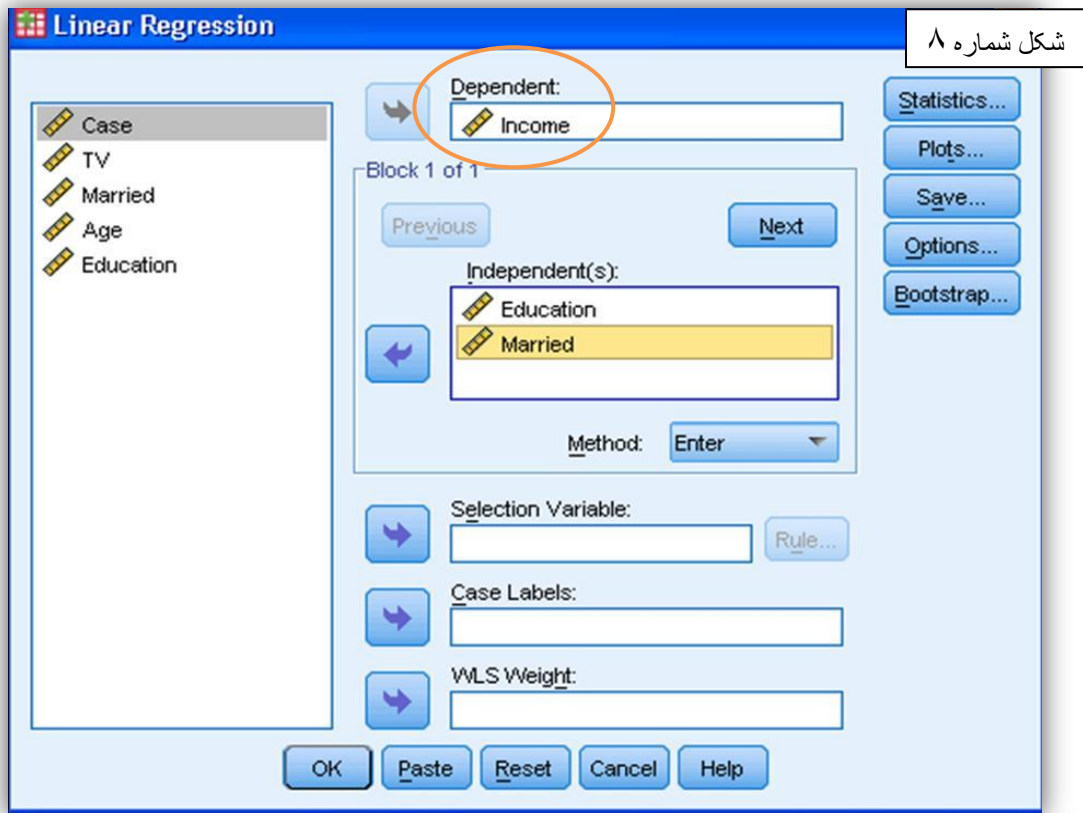
و مدل تجربی بدست آمده به صورت زیر است:



تحلیل نتایج:

از میان چهار متغیر مستقل که بر مبنای مدل نظری دارای تأثیرات مستقیم بر متغیر میزان تماشای تلویزیون بودند، تنها دو متغیر درآمد و تاهل دارای تأثیر مستقیم معناداری بر متغیر میزان تماشای تلویزیون هستند و دو متغیر دیگر پژوهش (تحصیلات و سن) تأثیر معناداری بر متغیر میزان تماشای تلویزیون ندارند ($P > 0.05$).

۳- در این مرحله متغیر درآمد (Income) را به عنوان متغیر وابسته در نظر می گیریم و تأثیر دو متغیر تحصیلات و تاهل را بر روی آن آزمون می کنیم (می خواهیم بتای مسیرهای P_{41} و P_{42} را بدست بیاوریم).



نکته: متغیر درآمد در مرحله ۳ نقش یک متغیر وابسته (میان) را دارد و می توان تاثیر متغیرهای مستقل را بر روی آن سنجید.

جدول شماره ۳ بیان گر ضرایب مسیر است و معناداری یا عدم معناداری و نیز میزان ضرایب مسیر را نشان می دهد. نتایج بیان گر آن است که متغیرهای تحصیلات و تاهل دارای تاثیر معناداری بر متغیر درآمد هستند و جهت هر دو ضریب نیز مثبت است.

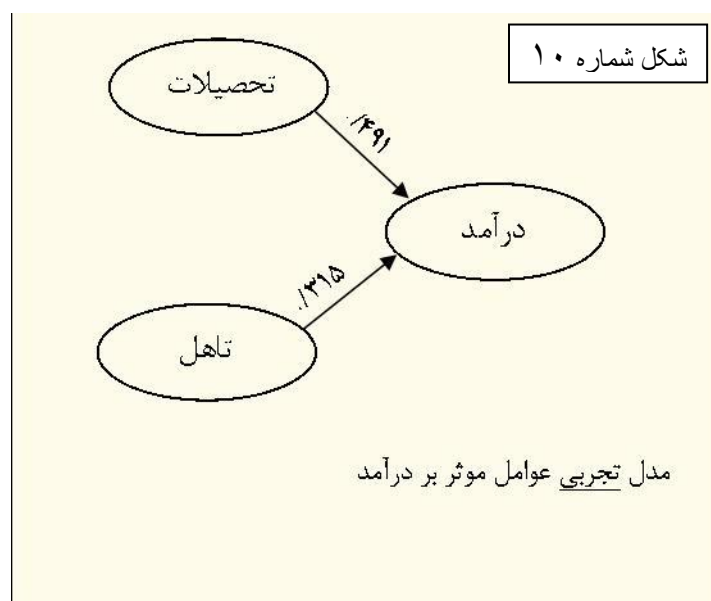
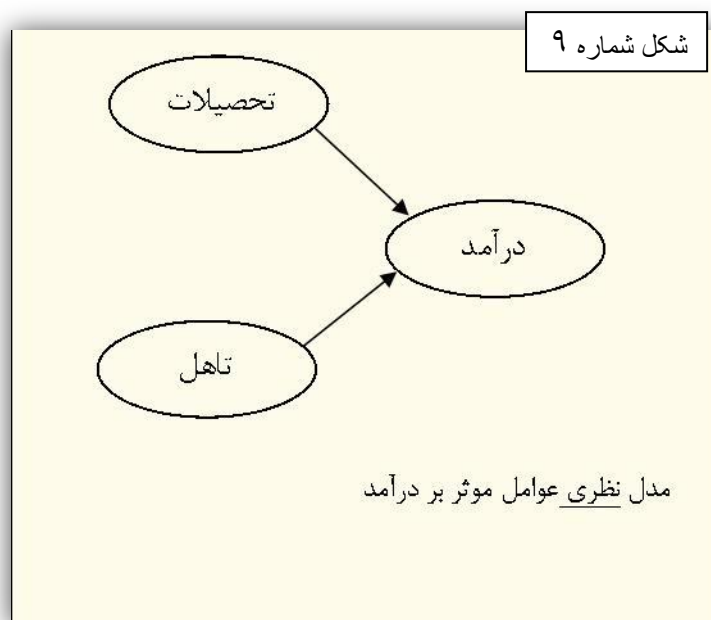
جدول شماره ۳

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	277.576	38.427		7.224	.000
	Education	28.851	4.324	.491	6.672	.000
	Married	164.678	38.479	.315	4.280	.000

a. Dependent Variable: Income

شکل شماره ۹ مدل مفهومی یا نظری عوامل موثر بر درآمد است. ما این مدل را نیز آزمون می کنیم. بر طبق نتایج به دست آمده از جدول شماره دو، تاثیر هر دو متغیر مستقل بر متغیر وابسته درآمد معنادار است و بنابراین مدل نظری با مدل تجربی انطباق کامل دارد.



تفسیر نتایج:

هدف از تحلیل رگرسیون برآورد ضرایب رگرسیونی (بتا) مسیرهای نمودار مسیر است. در نتیجه ما می بایست در پی شناسایی ضرایب رگرسیونی باشیم. و مدل نظری خود را با مدل تجربی به دست آمده تطبیق دهیم.

تاثیر متغیرهای مستقل بر متغیر وابسته نهایی به ۳ صورت است:

- (۱) صرفاً به صورت مستقیم
- (۲) صرفاً به صورت غیر مستقیم
- (۳) هم تاثیر مستقیم و هم غیر مستقیم

(۱) متغیرهایی که صرفاً به صورت مستقیم بر متغیر میزان تماشای تلویزیون تاثیر گذاشته اند:

متغیر سن و درآمد متغیرهایی هستند که تنها به صورت مستقیم بر متغیر میزان تماشای تلویزیون تاثیر دارند. نتایج جدول شماره ۱ نشان می دهد که متغیر درآمد دارای تاثیر معناداری بر متغیر وابسته است اما متغیر سن یا Age دارای تاثیر معناداری بر متغیر وابسته نیست ($P > .05$).

در نتیجه مسیر P_{53} را از مدل نظری حذف می کنیم.

(۲) متغیرهایی که صرفاً به صورت غیر مستقیم بر متغیر میزان تماشای تلویزیون تاثیر گذاشته اند:

در مثال ما متغیری که تنها به صورت غیر مستقیم بر متغیر وابسته نهایی تاثیر گذاشته باشد و تاثیر مسقیم بر متغیر وابسته نهایی نداشته باشد، وجود ندارد.

(۳) متغیرهایی که دارای تاثیر دوگانه (هم مستقیم و هم غیر مستقیم) بر متغیر میزان تماشای تلویزیون هستند:

دو متغیر تحصیلات و وضعیت تاهل، متغیرهایی هستند که دارای تاثیر دوگانه بر متغیر وابسته هستند. بر طبق نتایج جدول شماره ۱ از بین این دو متغیر وضعیت تاهل دارای تاثیر معناداری بر متغیر میزان تماشای تلویزیون است و میزان ضریب مسیر نیز برای این متغیر برابر $-.153$ می باشد. اما متغیر تحصیلات تاثیر معناداری بر متغیر وابسته ندارد ($P > .05$). در نتیجه ما مسیر شماره P_{51} را نیز از مدل حذف می کنیم.

نتایج جدول شماره ۲ نشان می دهد که در مجموع متغیرهای مستقل پژوهش توانسته اند حدود ۹۲ درصد از واریانس متغیر وابسته (میزان تماشای تلویزیون) را تبیین کنند.

نکته: در تحقیقات اجتماعی - اقتصادی معمولاً شناسایی کلیه عوامل تاثیرگذار بر متغیر وابسته برای محقق امکان پذیر نمی باشد. بنابراین متغیرهای تحلیل مسیر همواره می توانند تنها بخشی از واریانس متغیر وابسته را تبیین کنند. به همین دلیل در تحلیل مسیر آن چه که به عنوان اثر یا عوامل ناشناخته باقی می ماند به وسیله « e » که به « کمیت خطا » (Error term) معروف است، نشان داده می شود. مقدار e بیان گر میزان واریانس متغیری است که متغیرهای مستقل قبلی در مدل قادر به تبیین آن نبوده اند. با مجذور کردن e ، واریانس تبیین نشده به دست می آید. چنان چه مقدار e^2 از عدد ۱ کم شود مقدار واریانس تبیین شده (R^2) به دست می آید. به عبارت دیگر :

$$R^2 = 1 - e^2$$

در مثال ما مقدار e^2 برابر با ۰/۷۵/ شده است که نشان می دهد تنها ۷/۵ درصد از واریانس متغیر تماشای تلویزیون توسط مدل علی به دست آمده تبیین نمی شود.

$$e^2 = 1 - ۰/۹۲۵$$

در ادامه جهت تعیین مقادیر مسیرهای P_{41} و P_{42} از نتایج جدول شماره ۳ بهره می گیریم. بر طبق نتایج این جدول، متغیرهای تحصیلات و وضعیت تاهل هر دو دارای تاثیر معناداری بر متغیر درآمد هستند ($P < ۰/۰۵$).

گزارش نتایج:

نحوه محاسبه ضرایب مسیر مستقیم، غیرمستقیم و کل:

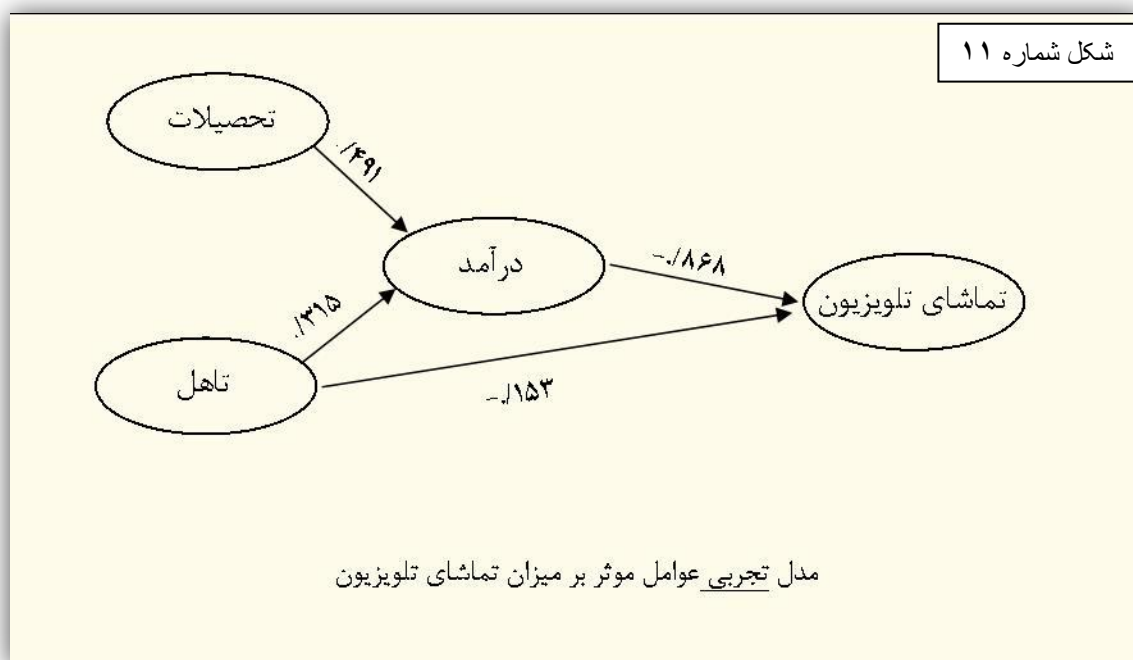
بعد از این که روش تحلیل مسیر را طی چندین مرحله اجرا کردیم، باید در قالب جدولی، نتایج مربوط به انواع تاثیرات مستقیم، غیر مستقیم و کل را نشان دهیم. جدولی مطابق زیر ترسیم می کنیم:

متغیرهای مستقل	انواع تاثیر		
	مستقیم	غیر مستقیم	کل (مجموع مستقیم و غیر مستقیم)
تحصیلات	—	-۰/۴۳	-۰/۴۳
تاهل	-۰/۱۵۳	-۰/۲۷	-۰/۴۲
سن	—	—	—
درآمد	-۰/۸۶۸	—	-۰/۸۶۸

اثرات غیر مستقیم هر متغیر برابر با حاصل ضرب مسیرهای جداگانه آن متغیر است. به عنوان مثال اثر غیر مستقیم متغیر تحصیلات برابر با حاصل ضرب مسیر های P_{41} و P_{54} است. یعنی حاصل ضرب ضرایب مسیر تحصیلات بر روی درآمد و درآمد بر روی تماشای تلویزیون $(-./۴۳ = -./۸۶۸ \times ./۴۹۱)$. همین طور تاثیر غیر مستقیم متغیر وضعیت تاهل بر تماشای تلویزیون برابر است با حاصل ضرب ضرایب مسیر تاهل بر روی درآمد و درآمد بر تماشای تلویزیون $(-./۲۷ = -./۸۶۸ \times ./۳۱۵)$.

تاثیر کل هر متغیر نیز برابر است با مجموع اثرات مستقیم و غیرمستقیم آن متغیر. در مورد متغیر تاهل اثر کل برابر شده است با $-.۴۲$ که مجموع اثرات مستقیم و غیر مستقیم این متغیر است $(-./۲۷ + -./۱۵۳)$.

مدل تجربی به دست آمده از تکنیک تحلیل مسیر در شکل شماره ۱۱ نشان داده شده است. همان طور که ملاحظه می کنیم در مدل تجربی به دست آمده دو مسیر P_{51} و P_{53} از مدل (دیاگرام) حذف شده اند.



منابع:

- حبیب پور، کرم و صفری، رضا (۱۳۸۸) راهنمای جامع کاربرد SPSS در تحقیقات پیمایشی، تهران: نشر لویه.
- کرلینجر و پدهازور (۱۳۶۶) رگرسیون چند متغیری در پژوهش رفتاری، ترجمه حسن سرایی، تهران: انتشارات مرکز نشر دانشگاهی.
- کلاتتری، خلیل (۱۳۸۲) پردازش و تحلیل داده ها در تحقیقات اجتماعی – اقتصادی، تهران: نشر شریف.
- منصورفر، کریم (۱۳۸۵) روش های پیشرفته آماری همراه با برنامه های کامپیوتری، تهران: انتشارات دانشگاه تهران.
- هومن، حیدرعلی (۱۳۸۴)، مدل یابی معادلات ساختاری با کاربرد نرم افزار لیزرل، تهران: انتشارات سمت.

مرکز خدمات آماری خوارزمی

انجام تحلیل آماری پایان نامه کارشناسی ارشد و دکترا و مقالات ISI

با نرم افزارهای SPSS – LISREL – AMOS – PLS – Eviews و شبکه های عصبی با Matlab

ایمیل: RKarimi777@yahoo.com

سایت: www.Kharazmi-Statistics.ir

www.SPSS100.ir

رامین کریمی: ۰۹۱۲۷۶۹۴۰۶۶

مؤلف کتاب "راهنمای آسان تحلیل آماری با SPSS"