

۷ کارنیل، بزرگترین شبکه موفقیت ایرانیان می باشد، که افرادی زیادی توانسته اند با آن به موفقیت برسند، فاطمه رتبه ۱۱ کنکور کارشناسی، محمد حسین رتبه ۶۸ کنکور کارشناسی، سپیده رتبه ۳ کنکور ارشد، مریم و همسرش راه اندازی تولیدی مانتو، امیر راه اندازی فروشگاه اینترنتی، کیوان پیوستن به تیم تراکتور سازی تبریز، میلاد پیوستن به تیم صبا، مهسا تحصیل در ایتالیا، و.... این موارد گوشه از افرادی بودند که با کارنیل به موفقیت رسیده اند، شما هم می توانید موفقیت خود را با کارنیل شروع کنید.

برای پیوستن به تیم کارنیلی های موفق روی لینک زیر کلیک کنید.

www.karnil.com

همچنین برای ورود به کانال تلگرام کارنیل روی لینک زیر کلیک کنید.

<https://telegram.me/karnil>



آمار پیشرفته در علوم اجتماعی

با کاربرد *lisrel* و *spss*

دکتر توکل آقاییاری هییر^۱

تهیه کننده: الناز اسمعیل پور^۲

۱۳۹۱

^۱ عضو هیئت علمی دانشگاه تبریز aghayarih@yahoo.com

^۲ دانشجوی کارشناسی ارشد دانشگاه تبریز elnaz.esmailpoor@ymail.com

فهرست:

۳مقدمه
۴کاربرد تحلیل عاملی در رگرسیون
۱۱کاربرد رگرسیون در علوم اجتماعی
۲۳کاربرد رگرسیون لجستیک در علوم اجتماعی
۳۳مدل‌های ساختار کواریانس در علوم اجتماعی
۴۴ضمایم

مقدمه:

نوشته حاضر با توجه به مباحث کلاس آمار پیشرفته در علوم اجتماعی، دکتر توکل آقایاری، گردآوری شده است و همچنین برای قابل درک شدن مباحث با توجه با داده‌های گردآوری شده در شهر تبریز (۱۳۸۹) مربوط به نگرش‌های خانواده، ازدواج و طلاق در نرم افزارهای SPSS اجرا شده است. در بخش آخر مدل لیزرل با توجه به این که مدل *converge* نشد و مشکل از نوع داده‌ها بود^۱ که عملاً برطرف شدنی نبود از داده‌های تحقیقی در مورد رفتار شهروندی استفاده شده است.

الناز اسمعیل پور

^۱ بیشتر متغیرهای مستقل با یک گویه سنجش شده‌اند که این باعث پایین آمدن درجه آزادی و *not converge* شدن مدل شد.

تحلیل عاملی

تحلیل عاملی factor analysis دو کاربرد عمده را شامل می‌شود:

۱. تقلیل داده‌ها (data reduction)

۲. تست پایایی و روایی تحقیق

پیش‌فرض‌های تحلیل عاملی:

۱. کفایت نمونه‌گیری (sampling adequacy): یعنی گویه‌ها از تمام فضای مفهومی انتخاب شوند. برای مثال در سنجش سرمایه اجتماعی باید از همه متغیرها که در مفهوم سرمایه اجتماعی وجود دارند چند گویه آورده شود.

آزمون KMO مشخص می‌کند که گویه‌ها تمام فضای مفهومی را دربرمی‌گیرند یا نه. این شاخص باید بیشتر از ۰/۷ باشد، البته در برخی نوشته‌ها بیش از ۰/۵ نیز گفته شده ولی در کل هرچه به یک نزدیک‌تر باشد بهتر است.

۲. وجود رابطه خطی متغیرها: که از طریق معنی‌داری آزمون Bartlett's می‌توان گفت تفکیک عامل‌ها درست انجام شده. فرضیه یک در اینجا این است که گویه‌ها با هم رابطه خطی دارند.

بار عاملی:

همبستگی تک تک متغیرها با عامل‌ها را نشان می‌دهد. علامت بار عاملی نشان می‌دهد متغیرهایی که استفاده شده‌اند به چه نحوی ترکیب شده‌اند، زمانی که یک گویه هم‌جهت نباشد علامت بار عاملی آن گویه منفی خواهد آمد. انتظار می‌رود علامت بارهای عاملی همگی یا مثبت و یا منفی باشند در غیر اینصورت باید گویه‌ها هم جهت شوند.

معمولا گفته می‌شود بارهای عاملی ۰/۴ و بیشتر میزان مناسبی می‌باشد. البته اگر تعداد نمونه خیلی کم باشد بارهای عاملی بالاتر ممکن است معنی‌دار شود و اگر تعداد نمونه خیلی زیاد بود، بارهای عاملی خیلی پائین‌تر نیز معنی‌دار می‌شوند. برای مثال در نمونه‌های بزرگ بار عاملی ۰/۳ کنار گذاشته می‌شود در این صورت اگر تئوری حمایت‌کننده‌ای دارید آن را برگردانید.

برحسب کاربرد تحلیل عاملی از دو نوع تحلیل عاملی می‌توان صحبت کرد:

۱. تحلیل عاملی اکتشافی (Exploratory Factor Analysis). ۲. تحلیل عاملی تأییدی (Conformatory Factor Analysis).

تحلیل عاملی اکتشافی

محقق بدون پیش فرض قبلی از نرم افزار می خواهد داده ها را بعدبندی کند. تلاش بر این است که عامل هایی استخراج شوند که حداکثر واریانس موجود در بین همه گویه ها را تبیین کنند. در شرایط ایده آل عامل هایی که استخراج می شوند صد در صد واریانس گویه ها را تبیین می کنند ولی معمولاً برخی گویه تبیین نمی شوند و آن را "واریانس تبیین نشده" و بخش تبیین شده را نیز "واریانس مشترک" می گویند. در تحلیل عاملی نیز اصل پارسی مونی سر جایش است هدف این است که بارهای عاملی کمتر با واریانس بیشتری را تبیین کنیم.

در تحلیل عاملی اکتشافی یکی از محورهای کار محقق نام گذاری عامل هاست. بعد از اینکه نرم افزار عامل بندی کرد باید براساس محتوای عامل ها (گویه هایی که در عامل ها وجود دارند) آنها را نام گذاری کرد. چراکه هدف استفاده از عامل ها در تحلیل های بعدی (برای مثال رگرسیون) می باشد.

چند عامل مناسب است؟

معمولاً عامل هایی مهم اند که بارهای عاملی بالایی دارند ولی برای تعیین تعداد عامل ها معمولاً چند شاخص وجود دارد:

۱. مقادیر ویژه (*Eigen Values*): یعنی مقدار واریانس موجود در متغیرها که توسط عامل ها تبیین می شود. مقادیر ویژه یک و بالاتر از یک را معمولاً عامل در نظر می گیرند. هر چه میزان مقادیر ویژه بیشتر باشد یعنی نقش آن عامل در تبیین واریانس بیشتر است.

۲. نمودار سنگ ریزه (*Scree Plot*): جایی که نمودار مربوطه به شکل افقی درمی آید همان جا cut point محسوب می شود.

زمانی که تعداد نمونه بزرگتر از ۲۵۰، واریانس مشترک (communalities) بزرگتر از ۰/۶ و تعداد گویه ها کمتر از ۳۰ باشد برای بعدبندی و تعیین تعداد عامل ها از شاخص مقادیر ویژه استفاده می شود و در نمونه های بالاتر از ۲۰۰، با میزان واریانس مشترک کمتر از ۰/۶ و تعداد گویه های بالاتر از ۳۰ از نمودار سنگ ریزه استفاده می شود.

چرخش عوامل

برحسب اینکه عواملی که استخراج می‌کنیم چه ویژگی‌هایی داشته باشند، اینکه اجازه همبسته بودن یا رابطه داشتن عامل‌ها را بدهیم یا نه، دو نوع عمده چرخش‌ها را خواهیم داشت:

۱. چرخش متعامد (*Ovthogona*): پیش‌فرض این نوع چرخش عدم رابطه بین عامل‌های استخراجی می‌باشد. اگر هدف استخراج عوامل مستقل از هم باشد، چرخش‌های متعامد بکار خواهد رفت. در تحقیقات علوم اجتماعی عمدتاً نمی‌شود فرض کرد که عوامل مستقل از هم باشند ولی با این حال چرخش متعامد زیاد بکار می‌رود و فرض‌اش نادیده گرفته می‌شود!
- سه نوع چرخش متعامد وجود دارد:

۱. *Varimax*: به حداکثر رساندن واریانس تبیین شده و طبیعتاً تولید عامل‌های زیاد.

۲. *Covaritimax*: به حداقل رساندن تعداد عامل‌هایی که برای تبیین واریان هر متغیر لازم است.

۳. *Equemax*: چرخشی منطقی‌تر برای تعدیل دو چرخش قبلی.

۲. چرخش متمایل (*Obique*): در این نوع چرخش‌ها پیش‌فرض این است که عامل‌های استخراجی مستقل از هم نیستند و بین عامل‌ها همبستگی وجود دارد.

Promax و *Direct bimin* از نوع چرخش‌های متمایل می‌باشند.

تحلیل عاملی تائیدی

محقق از قبل با فهم روشنی از عامل‌ها و اینکه کدام گویه مربوط به کدام عامل است وارد می‌شود. در اینجا عمدتاً هدف این است که چیزی که در مدل مورد نظر داشتیم با داده‌ها منطبق است یا نه؟ بنابراین در تحلیل عاملی تائیدی بدنبال تائید مدل نظری تحقیق هستیم.

در اینجا محقق از نرم‌افزار می‌خواهد فقط یک عامل بیاورد بعد سراغ تحلیل بارهای عاملی می‌رود، اینکه آیا تمام گویه‌ها در ساخت عامل تاثیر معنی‌داری داشته‌اند یا نه؟ معمولاً بارهای عاملی بالای ۰/۴ میزان خوبی است.

تحلیل عاملی برای تشخیص گویه‌ها نامناسب روش خوبی است. گویه‌ها با تبیین پائین میزان آلفای کرونباخ را بالا می‌برند.

نحوه اجرای تحلیل عاملی اکتشافی

برای اجراء تحلیل عاملی ابتدا به منوی Analysis/Data Reduction/Factor Analysis می‌روید.

به descriptive می‌روید و گزینه‌ها زیر را انتخاب می‌کنید:

Coefficients/ Significant Levels/ Determinant / KMO and Bartlett's of...

بعد در Rotation برحسب پیش‌فرض‌ها نوع چرخش را انتخاب می‌کنیم. و گزینه Scree Plots را نیز انتخاب می‌کنیم تا نمودار سنگ ریزه نیز بیاید.

در Extraction گزینه Covariance matrix را انتخاب می‌کنیم تا عامل‌ها را براساس کوراریانس داده‌ها ایجاد کند.

در scores اگر خواستید متغیر جدیدی ایجاد شود گزینه Save as Variable را انتخاب کنید. در اینصورت هر عامل به عنوان یک متغیر ذخیره خواهد شد. Display Factor Score... مشخص می‌کند عامل‌های استخراج شده چقدر همبسته‌اند.

در Options گزینه Sort by size باعث می‌شود ضرایب به ترتیب بیایند.

در selection Variable می‌توانید با وارد کردن یک متغیر تحلیل‌ها را به یک مورد خاص محدود کنید مثلاً برای زن و مرد جداگانه تحلیل کند.

نحوه اجرای تحلیل عاملی تأییدی

برای اجراء تحلیل عاملی تأییدی همانند تحلیل عاملی اکتشافی عمل می‌کنیم با این تفاوت که در Extraction در گزینه Number of factor تعداد عامل‌ها را یک می‌زنیم.

تحلیل عاملی تأییدی کاملاً براساس تئوری است اگر می‌خواهیم از بارهای عاملی به عنوان شاخص استفاده کنیم بهتر است منفی را از بین ببریم کوچکترین عدد شاخص را پیدا می‌کنیم در تحلیل عاملی تأییدی اگر KMO بالای ۰/۶، بارتلت تست معنی‌دار و بارهای عاملی بالای ۰/۴ باشند می‌توان گفت مدل اندازه‌گیری ما داده‌ها مطابقت دارند.

تفسیر جداول تحلیل عاملی:

در جدول شماره یک همانطور که گفته شد میزان KMO نشان دهنده اعتبار سازه‌ای تحقیق می‌باشد که در اینجا کمتر از ۰/۷ است ولی چون تقریباً نزدیک است می‌توان قبول کرد. همچنین آزمون بارتلتس نیز معنی دار است یعنی تفکیک عامل‌ها درست انجام شده است.

جدول شماره یک

KMO and Bartlett's Test ^a		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.623
Bartlett's Test of Sphericity	Approx. Chi-Square	176.467
	df	21
	Sig.	.000
a. Based on correlations		

در جدول شماره دو communalities بیانگر واریانسی است که هر یک از گویه‌ها تبیین می‌کنند. مقدار Initial به لحاظ تئوریک است و فرض این است که همه واریانس یک متغیر توسط یک گویه تبیین می‌شود. بنابراین همه گویه‌ها یک شده ولی در Extraction میزان واریانسی است که توسط عوامل استخراج شده تبیین می‌شود و هرچه به یک نزدیک باشد بهتر است.^۱

جدول شماره دو

	Communalities			
	Raw		Rescaled	
	Initial	Extraction	Initial	Extraction
talag-masaleei jedi dar jamee	.880	.015	1.000	.018
talag-padideei shom	2.568	1.725	1.000	.672
talag-talage tavafogi	2.349	1.877	1.000	.799
talag-taeene sharayete sakht baraye talag	1.955	.735	1.000	.376
talag-taahode daemee baraye zendegi	2.364	1.449	1.000	.613
talag-darkhaste talag dar har zaman va be har dalil	1.993	.895	1.000	.449
talag-tarbiate sahihe farzandane talag emkan pazir nist	1.850	.255	1.000	.138
Extraction Method: Principal Component Analysis.				

جدول شماره سه مقدار واریانس کلی تبیین شده را نشان می‌دهد. اگر ملاک تعیین تعداد عامل‌ها مقادیر ویژه باشد جدول شماره سه و اگر ملاک تعیین تعداد عامل‌ها نمودار سنگ ریزه باشد نمودار را گزارش می‌کنیم. در جدول شماره سه total نشان‌دهنده میزان مقادیر ویژه است که باید بالاتر از یک باشند تا به عنوان یک عامل حساب کنیم. معمولاً عامل‌های اول عامل‌های مهمتری‌اند.

^۱ به مقادیر rescaled توجه شود.

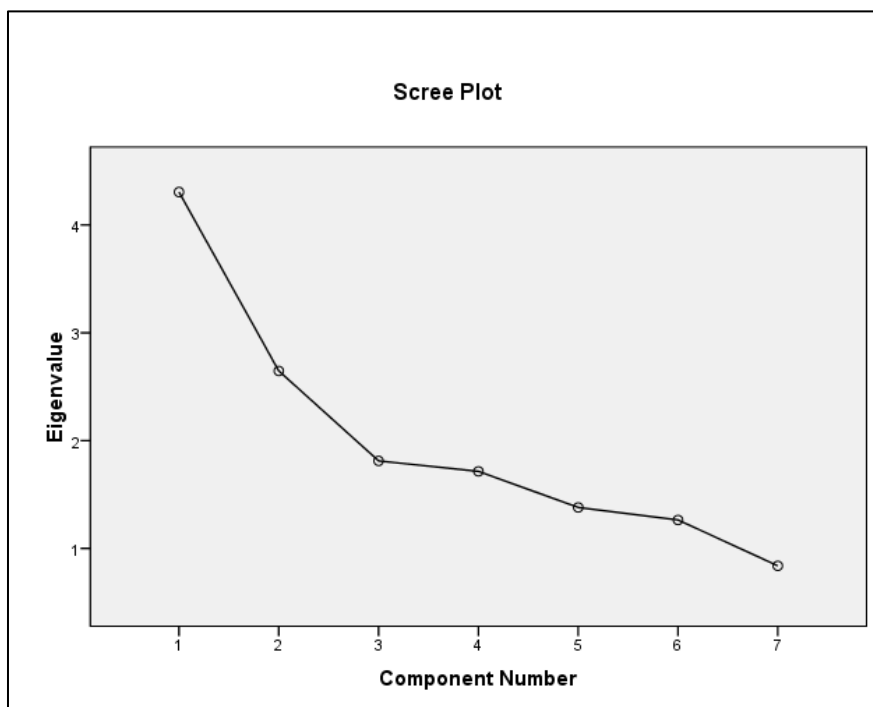
جدول شماره سه

Total Variance Explained										
	Component	Initial Eigenvalues ^a			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
		Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
Raw	1	4.305	30.839	30.839	4.305	30.839	30.839	4.296	30.774	30.774
	2	2.646	18.953	49.792	2.646	18.953	49.792	2.655	19.019	49.792
	3	1.812	12.977	62.769						
	4	1.715	12.283	75.052						
	5	1.380	9.887	84.940						
	6	1.264	9.055	93.994						
	7	.838	6.006	100.000						
Rescaled	1	4.305	30.839	30.839	1.877	26.814	26.814	1.866	26.662	26.662
	2	2.646	18.953	49.792	1.187	16.954	43.768	1.197	17.106	43.768
	3	1.812	12.977	62.769						
	4	1.715	12.283	75.052						
	5	1.380	9.887	84.940						
	6	1.264	9.055	93.994						
	7	.838	6.006	100.000						

Extraction Method: Principal Component Analysis.

a. When analyzing a covariance matrix, the initial eigenvalues are the same across the raw and rescaled solution.

نمودار سنگ ریزه



جدول شماره چهار عوامل چرخش نیافته است که تفسیرپذیری پایینی دارند

جدول شماره چهار

Component Matrix ^a				
	Raw		Rescaled	
	Component		Component	
	1	2	1	2
talag-padideei shom	1.306	-.140	.815	-.087
talag-taahode daemee baraye zendegi	1.202	-.063	.782	-.041
talag-taeene sharayete sakht baraye talag	.845	.141	.605	.101
talag-tarbiate sahihe farzandane talag emkan pazir nist	.489	.127	.359	.093
talag-masaleei jedi dar jamee	.118	-.040	.126	-.042
talag-talage tavafoqi	-.187	1.357	-.122	.886
talag-darkhaste talag dar har zaman va be har dalil	.390	.862	.276	.610

Extraction Method: Principal Component Analysis.
a. 2 components extracted.

با توجه به بارهای عاملی ۵ گویه اول مربوط به یک عامل و ۲ گویه بعدی برای عامل دوم می‌باشد. که در نام-گذاری آن عامل‌ها با توجه به گویه‌ها عامل اول را نگرش منفی به طلاق و عامل دوم را نگرش مثبت به طلاق نام‌گذاری کردیم.

جدول شماره پنجم

Rotated Component Matrix ^a				
	Raw		Rescaled	
	Component		Component	
	1	2	1	2
talag-padideei shom	1.313	-.043	.819	-.027
talag-taahode daemee baraye zendegi	1.203	.026	.783	.017
talag-taeene sharayete sakht baraye talag	.833	.203	.596	.145
talag-tarbiate sahihe farzandane talag emkan pazir nist	.478	.163	.351	.120
talag-masaleei jedi dar jamee	.121	-.031	.129	-.033
talag-talage tavafoqi	-.287	1.340	-.187	.874
talag-darkhaste talag dar har zaman va be har dalil	.325	.888	.230	.629

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 3 iterations.

جدول شماره ششم

Component Transformation Matrix		
Component	1	2
1	.997	.074
2	-.074	.997

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

رگرسیون

در آمار تحلیل رگرسیون یک نام جمعی برای فنونی است که برای الگو سازی و تحلیل داده های مربوط به متغیر وابسته و داده های مربوط به یک یا چند متغیر مستقل بکار می رود. بنابراین، در تحلیل رگرسیون ما نه تنها می-توانیم ارتباط دو یا چند متغیر با هم را از طریق محاسبه ضریب همبستگی محاسبه نماییم، بلکه می توانیم از این ضریب به عنوان پایه ای برای پیش بینی مقدار یک متغیر از متغیر دیگر استفاده کنیم و در اینجا است که راه رگرسیون از ضریب همبستگی جدا می شود.

بنابراین هدف از رگرسیون: ۱. برآورد تاثیر متغیرها ۲. پیش بینی

در رگرسیون باید اصل پارسی مونی (حداقل مطالعه، حداکثر نتیجه) را رعایت کرد. مدل هرچه ساده تر و قوی تر بهتر. متغیر کمتر، تبیین واریانس بیشتر.

فرض های زیر بنایی رگرسیون

۱. تعیین مدل مناسب

۲. متغیرهای مستقل باید به صورت خطی مستقل از هم باشند، یعنی شما فقط متغیرهای مستقل را که با متغیر وابسته رابطه دارند و با یکدیگر هیچ رابطه ای ندارند وارد معادله کنید. با استفاده از آزمون vif و تولرانس می-توان به خطی بودن رابطه متغیرها پی برد.

۳. خطی بودن رابطه ها: رابطه متغیرها را با خط مستقیم برآورد کنید که با آزمون f مشخص می شود

۴. عدم تاثیر بازگشتی مدل

۵. پایایی داده ها

۶. استقلال مشاهدات: با استفاده از آزمون T

۷. نبود مقادیر پرت (data explore): با استفاده از جدول Casewise Diagnostics غربالگری مورد به مورد.

۸. نرمال بودن داده ها: وقتی داده ها نرمال نبود:

۱. داده‌ها را شبه نرمال می‌کنیم (انتقال داده‌ها)

۲. استفاده از مدل جایگزین (ناپارامتریک)

روش انتقال داده‌ها: *چوله* به *راست*: جذر گیری، انتقال لگاریتمی، معکوس کردن مقادیر و ...

چوله به *چپ*: توان رساندن و ...

۹. ساختار رگرسیون مبتنی بر داده‌های کمی است و متغیرهای مستقل نیز می‌بایست به صورت فاصله‌ای یا نسبی اندازه گیری شوند. اما، استفاده از داده‌های اسمی یا رتبه‌ای برای متغیرهای مستقل منوط به رعایت شرایطی امکان‌پذیر خواهد بود.

اگر داده‌های ما از نوع مقوله‌ای باشند و بخواهیم از تحلیل رگرسیون خطی یا چند متغیری استفاده کنیم، ابتدا باید این داده‌ها را به صورت متغیرهای مجازی (*dummy variable*) که حالت دوجویی داشته باشند، تبدیل نماییم و در صفحه گستر *spss* در پنجره *data editor* هر مقوله را همانند یک متغیر مجزا به حساب آورده و وارد یک ستون مجزا نماییم.

محاسبه رگرسیون خطی یا چند متغیره با نرم افزار *spss*

برای محاسبه رگرسیون در *spss* ابتدا داده‌های مربوط به متغیر مستقل و وابسته خود را در ستون‌های مجزا در صفحه گستر *data editor* وارد کنید. بعد روی منوی *analyze* کلیک کنید و از پنجره *regression* را انتخاب نمایید و از پنجره مقابل آن *linear...* را انتخاب کنید تا پنجره *linear regression* پدیدار شود.

متغیر وابسته را در *dependent* و متغیرهای مستقل را در *independent* قرار می‌دهیم. بعد سراغ روش می‌رویم، در این قسمت ۵ روش وجود دارد: (۱) *enter*: کلیه متغیرهای مستقل به طور همزمان وارد تحلیل می‌شوند. (۲) *stepwise*: که در آن به ترتیب از قوی‌ترین متغیرها براساس میزان *F* یک به یک وارد معادله می‌شوند و این کار تا زمانی ادامه می‌یابد که خطای آزمون معناداری به ۵ درصد برسد. (۳) *remove*: که طی آن همه متغیرهای وارد شده به کادر در روش *enter* کنار گذاشته می‌شود. (۴) *backward*: که طی آن همه متغیرها وارد معادله می‌شوند و به مرور متغیرهای ضعیفتر حذف می‌شوند. این کار تا زمانی ادامه می‌یابد که

آمار پیشرفته در علوم اجتماعی

خطای معناداری آزمون به ۱۰ درصد برسد. ۵) forward: که در آن به ترتیب از قوی‌ترین متغیرها براساس بالاترین همبستگی با متغیر وابسته وارد معادله می‌شوند.

نتایج رگرسیون

اولین جدول رگرسیون درمورد آمار توصیفی متغیرهای مستقل و وابسته می‌باشد. که این جدول در گزارش محقق لازم نیست.

جدول شماره ۱: آمار توصیفی

Descriptive Statistics			
	Mean	Std. Deviation	N
Ntalag	3.1741	.98717	307
majazi2.mahalezendegi	.2280	.42024	307
majazi1.mahalezendegi	.3453	.47623	307
Zan	.6059	.48946	307
Mard	.3941	.48946	307
Paygaheejtemaee	5.2288	1.85917	307
SRelation	1.5415	1.00246	307
Sen	36.0717	15.53053	307

جدول شماره ۲ همبستگی همه متغیرها آورده شده است

جدول شماره ۲: همبستگی

Correlations									
		Ntala g	majazi2.mahalezendeg i	majazi1.mahalezendeg i	zan	Mard	paygaheejtemae e	SRelatio n	Sen
Pearson Correlatio n	Ntalag	1.000	-.086	.152	.132	-.132	.152	.251	-.278
	majazi2.mahalezendeg i	-.086	1.000	-.395	-.102	.102	-.099	.076	-.011
	majazi1.mahalezendeg i	.152	-.395	1.000	.039	-.039	.104	-.057	-.003
	zan	.132	-.102	.039	1.000	0	-.232	.089	-.008
	mard	-.132	.102	-.039	0	1.000	.232	-.089	.008
	paygaheejtemae e	.152	-.099	.104	-.232	.232	1.000	-.037	-.105
	Srelation	.251	.076	-.057	.089	-.089	-.037	1.000	-.393
	sen	-.278	-.011	-.003	-.008	.008	-.105	-.393	1.000
Sig. (1- tailed)	Ntalag	.	.066	.004	.010	.010	.004	.000	.000
	majazi2.mahalezendeg i	.066	.	.000	.037	.037	.042	.093	.427
	majazi1.mahalezendeg i	.004	.000	.	.248	.248	.034	.159	.480
	zan	.010	.037	.248	.	.000	.000	.059	.445
	mard	.010	.037	.248	.000	.	.000	.059	.445
	paygaheejtemae e	.004	.042	.034	.000	.000	.	.261	.033
	Srelation	.000	.093	.159	.059	.059	.261	.	.000
	sen	.000	.427	.480	.445	.445	.033	.000	.
N	Ntalag	307	307	307	307	307	307	307	307
	majazi2.mahalezendeg i	307	307	307	307	307	307	307	307
	majazi1.mahalezendeg i	307	307	307	307	307	307	307	307
	zan	307	307	307	307	307	307	307	307
	mard	307	307	307	307	307	307	307	307
	paygaheejtemae e	307	307	307	307	307	307	307	307
	Srelation	307	307	307	307	307	307	307	307
	sen	307	307	307	307	307	307	307	307

جدول شماره ۳ مربوط به متغیرهایی است که وارد معادله رگرسیونی شده‌اند و روش آنها نیز آورده شده که در گزارش محقق این جدول به درد نمی‌خورد.

جدول شماره ۳: متغیرهای وارد شده

Variables Entered/Removed ^a			
Model	Variables Entered	Variables Removed	Method
1	Sen	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	Srelation	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
3	majazi 1.mahalezendegi	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
4	Paygaheejtemaee	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
5	Mard	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
a. Dependent Variable: Ntag			

جدول شماره ۴ یکی از جداول مهم در گزارش‌ها محسوب می‌شود. در اینجا اگر روش شما enter باشد همه گام‌ها گزارش می‌شود و لی اگر از روش گام به گام (stepwise) استفاده کرده‌اید فقط گام آخر گزارش داده خواهد شد.

R₂ Square (ضریب تعیین): نشان دهنده درصدی از متغیر وابسته است که بوسیله متغیر مستقل وارد شده تبیین می‌شود. هرچه به صد نزدیک‌تر باشد بهتر است چون در این صورت مدل ما قدرت تبیین بالایی دارد. در این مدل ۵ متغیر مستقل وارد شده ۱۶/۲ درصد واریانس متغیر وابسته را تبیین می‌کند و بقیه توسط متغیرهای خارج از مدل که مدنظر محقق نیستند تبیین می‌شوند.

Adjusted R Square (ضریب تعیین تصحیح شده): در اینجا درجه آزادی هم حساب می‌شود و دقیق‌تر از ضریب تعیین می‌باشد لازم است که محقق در گزارش خود این ضریب را گزارش نماید. بنابراین در مثال ما ۱۴/۸ درصد واریانس نگرش افراد نسبت به طلاق توسط متغیرهای سن، روابط اجتماعی، تعلق به طبقه بالا، پایگاه اجتماعی و مرد بودن تبیین می‌شود.

Std. error of the estimate: مقدار خطای برآورد

آمار پیشرفته در علوم اجتماعی

F changing: نشانگر سطح اهمیت هر متغیر است که با اضافه شدن گام به گام هر متغیر قابل مشاهده است ولی در گزارش محقق نمی آید.

Durbin Watson: نشانگر این است که آیا نمونه ما نمونه نمایایی است یا نه؟ که باید بین $2/5$ _ $1/5$ باشد. در مثال ما دوربین واتسون $1/95$ می باشد که عددی قابل قبول است و این یعنی خطاها از هم مستقل و ما مجاز به انجام تحلیل رگرسیونی هستیم.

جدول شماره ۴: model summary

Model Summary ^f										
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Change Statistics					Durbin-Watson
					R Square Change	F Change	df1	df2	Sig. Change	
1	.278 ^a	.077	.074	.94983	.077	25.533	1	305	.000	
2	.318 ^b	.101	.095	.93902	.024	8.062	1	304	.005	
3	.356 ^c	.127	.118	.92693	.026	8.987	1	303	.003	
4	.377 ^d	.142	.130	.92052	.015	5.234	1	302	.023	
5	.402 ^e	.162	.148	.91131	.020	7.133	1	301	.008	1.956
a. Predictors: (Constant), sen										
b. Predictors: (Constant), sen, SRelation										
c. Predictors: (Constant), sen, SRelation, majazi1.mahalezendegi										
d. Predictors: (Constant), sen, SRelation, majazi1.mahalezendegi, paygaheejtemaee										
e. Predictors: (Constant), sen, SRelation, majazi1.mahalezendegi, paygaheejtemaee, mard										
f. Dependent Variable: Ntalag										

آمار پیشرفته در علوم اجتماعی

جدول شماره ۵ نشان می‌دهد که آیا مدل شما مدل معناداری هست یا نه؟ در اینجا معناداری F باید کمتر از $0/05$ باشد در این صورت متغیر مستقل ترکیب خطی و مستقلاً ب متغیر وابسته دارد. چون روش ما گام به گام است فقط معناداری گام آخر مهم است و آوردن آن در گزارش توسط محقق ضروری است. در مثال ما نیز مقدار F معنی دار می‌باشد.

جدول شماره ۵: Anova

ANOVA ^f						
Model		Sum of Squares	Df	Mean Square	F	Sig.
1	Regression	23.036	1	23.036	25.533	.000 ^a
	Residual	275.165	305	.902		
	Total	298.200	306			
2	Regression	30.144	2	15.072	17.093	.000 ^b
	Residual	268.056	304	.882		
	Total	298.200	306			
3	Regression	37.866	3	12.622	14.690	.000 ^c
	Residual	260.335	303	.859		
	Total	298.200	306			
4	Regression	42.301	4	10.575	12.480	.000 ^d
	Residual	255.899	302	.847		
	Total	298.200	306			
5	Regression	48.225	5	9.645	11.614	.000 ^e
	Residual	249.976	301	.830		
	Total	298.200	306			
a. Predictors: (Constant), sen						
b. Predictors: (Constant), sen, SRelation						
c. Predictors: (Constant), sen, SRelation, majazi1.mahalezendegi						
d. Predictors: (Constant), sen, SRelation, majazi1.mahalezendegi, paygaheejtemaee						
e. Predictors: (Constant), sen, SRelation, majazi1.mahalezendegi, paygaheejtemaee, mard						
f. Dependent Variable: Ntalag						

جدول شماره ۶ ضرایب B و $Beta$ را به ما نشان می‌دهد که در گزارش تحلیل رگرسیونی این جدول از جداول مهم می‌باشد.

ضرایب b و ثابت برای پیش‌بینی مرد استفاده قرار می‌گیرد و نشان‌دهنده میزان تاثیر هر یک از متغیرهای مستقل بر روی متغیر وابسته است. علامت منفی نشان دهنده تاثیر منفی متغیر مستقل بر متغیر وابسته می‌باشد. در مثال ما؛ با افزایش سن، نگرش مثبت به طلاق $0/012$ واحد کاهش می‌یابد. با افزایش روابط اجتماعی $0/175$ واحد نگرش مثبت به طلاق افزایش می‌یابد. تعلق به منطقه بالا نسبت به تعلق به منطقه متوسط $0/288$ واحد نگرش مثبت به طلاق را افزایش می‌دهد. با بالا رفتن پایگاه اجتماعی نگرش مثبت به طلاق $0/084$ واحد افزایش می‌یابد. مرد بودن نسبت به زن بودن $0/294$ نگرش مثبت به طلاق را کاهش می‌دهد.

Beta: ضریب رگرسیونی استاندارد شده است که متغیرهای اندازه گیری شده مختلف را براساس یک معیار استاندارد کرده و اهمیت نسبی هر متغیر مستقل را نشان می‌دهد و بدین ترتیب امکان مقایسه متغیرهای مستقل را فراهم می‌نماید. متغیری که بیشترین بتا را دارد بیشترین تاثیر را بر متغیر وابسته دارد به عبارتی بتا سهم هر متغیر را در تبیین متغیر وابسته مشخص می‌نماید.

در مثال ذکر شده تاثیر متغیر سن بر نگرش افراد نسبت به طلاق بیشتر از سایر متغیرهاست.

T: سطح معناداری آزمون t باید کمتر از 0.05 باشد در این صورت می‌توان گفت رابطه بین متغیرهای مستقل و وابسته معنی‌دار است.

Confidence interval for b (حدود اطمینان): حدود اطمینان برای ضرایب b را نشان می‌دهد.

Zero _ order (ضرایب همبستگی مرتبه صفر):

Partial correlation: فقط همبستگی متغیرهای داده شده را می‌سنجد و تاثیر متغیرهای دیگر را بر روی متغیر وابسته و مستقل کنترل می‌کند. برای مثال در رابطه بین X_1 و X_2 ، هم تاثیر X_3 و هم تاثیر متغیر Z را به عنوان یک متغیر مداخله‌گر کنترل می‌کند.

Part correlation: همبستگی متغیرهای داده شده را می‌سنجد ولی فقط تاثیر حذف شده را بر روی متغیر مستقل کنترل می‌کند. برای مثال در رابطه بین X_1 و X_2 ، فقط تاثیر X_3 را کنترل می‌کند.

Collinearity: زمانی که نیاز به بررسی دقیق وجود دارد مهم است و معمولاً در گزارش محقق آورده می‌شود. که شامل دو آمار ضریب تحمل و vif می‌باشد.

Tolerance (ضریب تحمل): باید بالا باشد، هرچه به صفر نزدیک باشد مشکل هم‌خطی متغیرهای مستقل وجود دارد.

Vif: vif بزرگتر و یا مساوی 10 نشان دهنده جدی بودن مساله هم خطی است و $vif = 1$ بدین معنی است که پیش‌بینی‌ها ناهمبسته اند یعنی بین متغیرهای مستقل وارد شده در مدل رگرسیونی مشکل هم‌خطی وجود ندارد.

جدول شماره ۶: ضرایب

Coefficients ^a											
Model		Unstandardized Coefficients		Standardized Coefficients	T	Sig.	Correlations			Collinearity Statistics	
		B	Std. Error	Beta			Zero-order	Partial	Part	Tolerance	VIF
1	(Constant)	3.811	.137		27.765	.000					
	sen	-.018	.003	-.278	-5.053	.000	-.278	-.278	-.278	1.000	1.000
2	(Constant)	3.405	.197		17.264	.000					
	sen	-.013	.004	-.212	-3.582	.000	-.278	-.201	-.195	.845	1.183
	SRelation	.165	.058	.168	2.839	.005	.251	.161	.154	.845	1.183
3	(Constant)	3.262	.200		16.269	.000					
	sen	-.013	.004	-.207	-3.545	.000	-.278	-.200	-.190	.845	1.184
	SRelation	.176	.058	.179	3.061	.002	.251	.173	.164	.842	1.188
	majazi1.mahalezendegi	.334	.111	.161	2.998	.003	.152	.170	.161	.996	1.004
4	(Constant)	2.872	.262		10.956	.000					
	sen	-.012	.004	-.190	-3.247	.001	-.278	-.184	-.173	.831	1.204
	SRelation	.187	.057	.190	3.254	.001	.251	.184	.173	.837	1.195
	majazi1.mahalezendegi	.309	.111	.149	2.776	.006	.152	.158	.148	.986	1.014
	paygaheejtemaee	.066	.029	.124	2.288	.023	.152	.131	.122	.972	1.029
5	(Constant)	2.918	.260		11.222	.000					
	sen	-.012	.004	-.190	-3.283	.001	-.278	-.186	-.173	.831	1.204
	SRelation	.175	.057	.177	3.062	.002	.251	.174	.162	.831	1.203
	majazi1.mahalezendegi	.288	.110	.139	2.610	.010	.152	.149	.138	.981	1.019
	paygaheejtemaee	.084	.029	.158	2.871	.004	.152	.163	.151	.919	1.088
	mard	-.294	.110	-.146	-2.671	.008	-.132	-.152	-.141	.935	1.069

a. Dependent Variable: Ntalag

جدول شماره ۷ مربوط به ضرایب بتا و دیگر ضرایب متغیرهایی است که از معادله خارج شده‌اند. این جدول در گزارش محقق نمی‌آید.

متغیرهایی که Partial Correlation بالایی دارند متغیر خوبی برای برگشتن به مدل محسوب می‌شوند.

جدول شماره ۷: متغیرهای خارج شده

Excluded Variables ^f								
Model	Beta In	t	Sig.	Partial Correlation	Collinearity Statistics			
					Tolerance	VIF	Minimum Tolerance	
1	majazi2.mahalezendegi	-.089 ^a	-1.620	.106	-.093	1.000	1.000	1.000
	majazi1.mahalezendegi	.151 ^a	2.771	.006	.157	1.000	1.000	1.000
	zan	.130 ^a	2.376	.018	.135	1.000	1.000	1.000
	mard	-.130 ^a	-2.376	.018	-.135	1.000	1.000	1.000
	paygaheejtemaee	.124 ^a	2.264	.024	.129	.989	1.011	.989
	SRelation	.168 ^a	2.839	.005	.161	.845	1.183	.845
2	majazi2.mahalezendegi	-.102 ^b	-1.869	.063	-.107	.994	1.006	.840
	majazi1.mahalezendegi	.161 ^b	2.998	.003	.170	.996	1.004	.842
	zan	.116 ^b	2.141	.033	.122	.991	1.009	.838
	mard	-.116 ^b	-2.141	.033	-.122	.991	1.009	.838
	paygaheejtemaee	.139 ^b	2.549	.011	.145	.982	1.019	.831
3	majazi2.mahalezendegi	-.045 ^c	-.772	.441	-.044	.841	1.189	.839
	zan	.109 ^c	2.033	.043	.116	.989	1.011	.834
	mard	-.109 ^c	-2.033	.043	-.116	.989	1.011	.834
	paygaheejtemaee	.124 ^c	2.288	.023	.131	.972	1.029	.831
4	majazi2.mahalezendegi	-.037 ^d	-.640	.523	-.037	.838	1.193	.831
	zan	.146 ^d	2.671	.008	.152	.935	1.069	.831
	mard	-.146 ^d	-2.671	.008	-.152	.935	1.069	.831
5	majazi2.mahalezendegi	-.019 ^e	-.333	.739	-.019	.827	1.210	.827
	zan	. ^e000	.	.000
a. Predictors in the Model: (Constant), sen								
b. Predictors in the Model: (Constant), sen, SRelation								
c. Predictors in the Model: (Constant), sen, SRelation, majazi1.mahalezendegi								
d. Predictors in the Model: (Constant), sen, SRelation, majazi1.mahalezendegi, paygaheejtemaee								
e. Predictors in the Model: (Constant), sen, SRelation, majazi1.mahalezendegi, paygaheejtemaee, mard								
f. Dependent Variable: Ntag								

جدول شماره ۸ همبستگی بین متغیرهای مستقل وارد شده را نشان می‌دهد، کوواریانس نیز در اینجا ذکر شده که کوواریانس صفر نشان دهنده عدم همبستگی بین متغیرهای مستقل و خوب است. این جدول نیز در گزارش محقق نمی‌آید.

جدول شماره ۸: ضرایب همبستگی

Coefficient Correlations ^a							
Model		sen	SRelation	majazi1.mahalezendegi	paygaheejtemaee	Mard	
1	Correlations	sen	1.000				
	Covariances	sen	1.222E-5				
2	Correlations	sen	1.000	.393			
		SRelation	.393	1.000			
	Covariances	sen	1.413E-5	8.613E-5			
		SRelation	8.613E-5	.003			
3	Correlations	sen	1.000	.394	.028		
		SRelation	.394	1.000	.064		
		majazi1.mahalezendegi	.028	.064	1.000		
	Covariances	sen	1.378E-5	8.430E-5	1.147E-5		
		SRelation	8.430E-5	.003	.000		
		majazi1.mahalezendegi	1.147E-5	.000	.012		
4	Correlations	sen	1.000	.400	.015	.128	
		SRelation	.400	1.000	.055	.079	
		majazi1.mahalezendegi	.015	.055	1.000	-.100	
		paygaheejtemaee	.128	.079	-.100	1.000	
	Covariances	sen	1.382E-5	8.530E-5	6.049E-6	1.365E-5	
		SRelation	8.530E-5	.003	.000	.000	
		majazi1.mahalezendegi	6.049E-6	.000	.012	.000	
		paygaheejtemaee	1.365E-5	.000	.000	.001	
5	Correlations	sen	1.000	.399	.015	.124	.001
		SRelation	.399	1.000	.060	.058	.080
		majazi1.mahalezendegi	.015	.060	1.000	-.113	.070
		paygaheejtemaee	.124	.058	-.113	1.000	-.233
		mard	.001	.080	.070	-.233	1.000
	Covariances	sen	1.354E-5	8.362E-5	5.957E-6	1.335E-5	3.980E-7
		SRelation	8.362E-5	.003	.000	9.687E-5	.001
		majazi1.mahalezendegi	5.957E-6	.000	.012	.000	.001
		paygaheejtemaee	1.335E-5	9.687E-5	.000	.001	.000
		mard	3.980E-7	.001	.001	.000	.012

a. Dependent Variable: Ntalag

جدول شماره ۹ یکی از راه‌های پیش‌بینی همخطی مدل می‌باشد.

جدول شماره ۹: جدول تعیین همخطی

Collinearity Diagnostics ^a									
Model	Dimension	Eigenvalue	Condition Index	Variance Proportions					
				(Constant)	sen	SRelation	majazi1.mahalezendegi	paygaheejtemaee	Mard
1	1	1.919	1.000	.04	.04				
	2	.081	4.859	.96	.96				
2	1	2.633	1.000	.01	.02	.03			
	2	.321	2.863	.00	.16	.47			
	3	.046	7.545	.99	.83	.50			
3	1	3.049	1.000	.01	.01	.02	.04		
	2	.591	2.271	.00	.01	.05	.90		
	3	.316	3.108	.00	.18	.43	.04		
	4	.045	8.246	.98	.81	.50	.03		
4	1	3.924	1.000	.00	.01	.01	.02	.01	
	2	.600	2.557	.00	.00	.04	.92	.00	
	3	.320	3.502	.00	.14	.47	.03	.01	
	4	.127	5.564	.00	.31	.15	.02	.53	
	5	.029	11.626	.99	.54	.33	.01	.46	
5	1	4.369	1.000	.00	.01	.01	.02	.00	.02
	2	.660	2.573	.00	.00	.00	.63	.00	.30
	3	.517	2.906	.00	.00	.12	.28	.00	.51
	4	.306	3.782	.00	.18	.38	.05	.00	.09
	5	.119	6.067	.00	.27	.16	.03	.56	.08
	6	.029	12.268	.99	.54	.33	.01	.43	.00

a. Dependent Variable: Ntalag

جدول غربالگری مورد به مورد مربوط به داده‌های پرت است و موردی که داده‌های آن پرت است را مشخص می‌سازد.

جدول شماره ۱۰: غربالگری مورد به مورد

Casewise Diagnostics ^a				
Case Number	Std. Residual	Ntalag	Predicted Value	Residual
205	-3.539	.62	3.8413	-3.22468

a. Dependent Variable: Ntalag

آمار پیشرفته در علوم اجتماعی

آخرین جدول مربوط به آمار باقیماندهها (تفاوت بین مقادیر پیش‌بینی شده و واقعی) می‌باشد(جدول شماره ۱۱).

جدول شماره ۱۱: آمار باقیماندهها

Residuals Statistics ^a					
	Minimum	Maximum	Mean	Std. Deviation	N
Predicted Value	2.0502	4.0849	3.1741	.39698	307
Std. Predicted Value	-2.831	2.294	.000	1.000	307
Standard Error of Predicted Value	.079	.218	.125	.024	307
Adjusted Predicted Value	1.9940	4.1018	3.1743	.39724	307
Residual	-3.22468	2.70368	.00000	.90383	307
Std. Residual	-3.539	2.967	.000	.992	307
Stud. Residual	-3.571	2.999	.000	1.002	307
Deleted Residual	-3.28496	2.76273	-.00012	.92221	307
Stud. Deleted Residual	-3.644	3.040	.000	1.006	307
Mahal. Distance	1.304	16.513	4.984	2.456	307
Cook's Distance	.000	.040	.003	.005	307
Centered Leverage Value	.004	.054	.016	.008	307

a. Dependent Variable: Ntalog

رگرسیون لجستیک (Binary logistic regression)

در مباحث مربوط به رگرسیون اغلب بحث برسر یافتن رابطه‌ی بین متغیر وابسته (Y) و مجموعه‌ای از متغیرهای مستقل است. در بسیاری از پژوهش‌ها متغیر وابسته دوحالتی است یعنی پاسخ‌ها تنها شامل دو حالت؛ وجود یا عدم وجود رابطه است یعنی متغیر وابسته ما اسمی می‌باشد که در این صورت از رگرسیون لجستیک استفاده می‌کنیم. لجستیک احتمال صفر یا یک شدن رابطه یا همان عدم وجود و وجود رابطه را نشان می‌دهد و بر پایه‌ی کی‌دو می‌باشد.

نسبت برتری (Odds Ratio): در عمل نسبتی است بین احتمال تجربه یک واقعه تقسیم بر عدم تجربه آن. بازده احتمالی در اینجا ۱-۰ می‌باشد اگر کمتر از ۰/۵ شد احتمال وقوع وجود ندارد اگر بیشتر از ۰/۵ شد احتمال وقوع وجود دارد.

مثال: احتمال ابتلا به سرطان ریه را بررسی می‌کنیم احتمال ابتلا به سرطان یک متغیر دو وجهی که کد ۱: مبتلا و کد ۰: سالم یا عدم ابتلا می‌باشد. از ۵۰ نفر ۱۰ نفر مبتلا و ۴۰ نفر سالم‌اند.

$$\text{احتمال ابتلا: } \frac{10}{50} \quad \text{احتمال سالم ماندن: } \frac{40}{50}$$

$$\text{نسبت برتری ابتلا شدن به سالم ماندن: } \frac{\frac{10}{50}}{\frac{40}{50}} = 0/25$$

حکم کلی: این واقعه به وقوع نمی‌پیوندد چون کمتر از ۰/۵ می‌باشد.

حال متغیر سیگار کشیدن / نکشیدن را به عنوان یک متغیر اسمی تاثیرگذار وارد می‌کنیم برای اینکه بدانیم سیگار چه تاثیری بر ابتلا دارد نسبت برتری آن به صورت زیر می‌آید:

نسبت برتری ابتلا به سرطان برای سیگاری‌ها

نسبت برتری ابتلا به سرطان برای غیر سیگاری‌ها

اگر یک شد سیگار در ابتلا به سرطان نقشی ندارد. اگر ۱/۲ شد احتمال ابتلا به سرطان برای سیگاری‌ها ۱/۲ برابر غیر سیگاری‌هاست. به عبارتی اگر بالاتر از یک شد احتمال ابتلا بالاتر و اگر کمتر از یک شد احتمال ابتلا کمتر می‌شود.

احتمال ابتلا برای سیگاری‌ها: $\frac{0/25}{0/75}$ احتمال ابتلا برای غیر سیگاری: $\frac{0/10}{0/90}$

نسبت برتری (EXP(b)): $\frac{\frac{0/25}{0/75}}{\frac{0/10}{0/90}}$

بنابراین احتمال ابتلا برای سیگاری‌ها ۳ برابر غیر سیگاری‌هاست.

حال می‌رسیم به پیش‌بینی که باید جایگاه افراد را بدرستی مشخص کنیم. پیش‌بینی در لجستیک بدین معناست که آیا می‌شود با استفاده از یکسری متغیرها جایگاه و... موقعیت افراد را بدرستی پیش‌بینی کرد؟ متغیر پیش‌بین (مستقل) جایگاه افراد را نشان می‌دهند این اطلاعات در SPSS در جدولی تحت عنوان با نام classification table می‌آید.

نحوه اجراء رگرسیون لجستیک

از مسیر زیر به کادر محاوره رگرسیون لجستیک می‌رویم:

Analyze/ regression/ binary logistic

در این کادر محاوره متغیر وابسته را به قسمت dependent منتقل می‌کنیم و متغیرهای مستقل را به قاب covariates منتقل می‌کنیم. گزینه categorical برای مرتب کردن متغیرهای فاصله‌ای مستقل می‌باشد. متغیرهای فاصله‌ای متغیرهای خوبی نیستند، یک واحد تغییر وقتی واحدها ریزند تاثیر معنی‌دار بوجود نمی‌آورد به سختی می‌شود تاثیر معنادار متغیرهای فاصله‌ای را در رگرسیون لجستیک پیدا کرد.

در reference category بهتر است first را علامت بزنی تا گزینه اول حساب شود.

از گزینه save می‌توانید مقادیر پیش‌بینی شده را بدست آورید. همچنین روی مقادیر باقیمانده نیز مدیریت کنید.

با کلیک روی option گزینه classification plot (احتمال وقوع یا عدم وقوع)، hosmer leme show goodness-of-fit (شاخص برآزش مدل)، classification on cut off (احتمال وقوع یا عدم وقوع ۰/۵) می‌باشند. و همچنین maximum iterations برای این است که اگر مدل‌ها خیلی پیچیده باشد و برآوردهای با ثبات نتواند ایجاد کند یک فرصت دیگر می‌دهد.

نتایج رگرسیون لجستیک:

جدول شماره یک داده‌های مربوط به missing مهم می‌باشد که برای خود محقق مهم است و در گزارش‌ها نمی‌آید.

جدول شماره یک

Case Processing Summary			
Unweighted Cases ^a		N	Percent
Selected Cases	Included in Analysis	313	92.1
	Missing Cases	27	7.9
	Total	340	100.0
Unselected Cases		0	.0
Total		340	100.0

a. If weight is in effect, see classification table for the total number of cases.

جدول شماره دو مربوط به کدهایی است که به متغیر وابسته داده‌ایم برای مثال در اینجا متغیر وابسته ما نگرش افراد نسبت به طلاق است که به نگرش مثبت به طلاق (احتمال وجود نگرش مثبت) کد ۱ و نگرش منفی به طلاق (عدم احتمال وجود نگرش مثبت) کد ۰ داده شده است.

جدول شماره دو

Dependent Variable Encoding	
Original Value	Internal Value
Non-probability	0
Probability	1

Block 0 مدل پایه می‌باشد و هیچکدام از متغیرها وارد مدل نشده‌اند. یکسری برآزش‌ها از متغیر وابسته است که در مراحل بعدی متغیرهای مستقل وارد و تفسیر می‌شوند.

Block 0: Beginning Block

جدول شماره یک (iteration history) می‌آید که میزان minis two log likelihood مهم است این همان ناتوانی‌های مدل است جاهایی که مدل خوب کار نمی‌کند. این میزان هرچه پایین تر باشد بهتر است

آمار پیشرفته در علوم اجتماعی

و با مدل نظری انطباق بیشتری دارد. جاهایی که صحبت از مقایسه داده‌ها می‌شود از میزان بهبودی صحبت می‌شود ولی در اینجا چون هنوز متغیرهای مستقل وارد نشده قابل تفسیر نیست.

جدول شماره سه

Iteration History ^{a,b,c}			
Iteration		-2 Log likelihood	Coefficients
			Constant
Step 0	1	431.219	.185
	2	431.219	.186
a. Constant is included in the model.			
b. Initial -2 Log Likelihood: 431.219			
c. Estimation terminated at iteration number 2 because parameter estimates changed by less than .001.			

جدول شماره چهار یکسری تفاسیر عمومی است و انتظار می‌رود در مراحل بعدی درصد کل (overall percentage) بالاتر برود. انتظار می‌رود احتمال وجود متغیر که در اینجا صد است چون صد در صد نگرش مثبت می‌باشد، در مراحل بعدی کمتر شود.

جدول شماره چهار

Classification Table ^{a,b}					
		Observed	Predicted		
			TALAG		Percentage Correct
			Non-probability	Probability	
Step 0	TALAG	Non-probability	0	142	.0
		Probability	0	171	100.0
	Overall Percentage				54.6
a. Constant is included in the model.					
b. The cut value is .500					

جدول شماره پنجم میزان $\text{Exp}(b)$ مهم است و در گزارش‌ها می‌آید ولی در مدل‌های بعدی این مدل پایه است و قابل تفسیر نیست. همچنین wald همان آزمون t است که باید به معنی‌داری آن توجه کرد اگر معنی‌دار بود سراغ تفسیر $\text{Exp}(b)$ می‌رویم.

جدول شماره پنجم

Variables in the Equation							
		B	S.E.	Wald	Df	Sig.	Exp(B)
Step 0	Constant	.186	.114	2.679	1	.102	1.204

در جدول شماره ششم همه متغیرهای مستقل آورده شده است.

جدول شماره ششم

Variables not in the Equation					
			Score	df	Sig.
Step 0	Variables	vazeatfaaleat	6.906	1	.009
		jens	3.232	1	.072
		tagiremahlezendegi	.392	1	.531
		tahsilat	14.558	1	.000
		vaziytetaahol	5.048	1	.025
	Overall Statistics			22.962	5

در Block 1 چون روش ما Enter می‌باشد همه متغیرهای مستقل یکجا وارد شده‌اند.

Block 1: Method = Enter

در جدول شماره هفتم با توجه به مدل صفر میزان minis two log likelihood کمتر شده و این نشان می‌دهد که متغیرهای مستقل مدل را بهبود داده‌اند. همچنین مدل در چرخش چهارم ثابت (converge) شده است.

جدول شماره هفتم

Iteration History ^{a,b,c,d}								
Iteration		-2 Log likelihood	Coefficients					
			Constant	vazeatfaaleat	jens	tagiremahlezendegi	tahsilat	Vaziytetaahol
Step 1	1	407.806	-1.925	-.522	.288	-.020	1.031	.049
	2	407.724	-2.067	-.560	.315	-.022	1.093	.057
	3	407.724	-2.069	-.560	.316	-.022	1.094	.057
	4	407.724	-2.069	-.560	.316	-.022	1.094	.057
a. Method: Enter								
b. Constant is included in the model.								
c. Initial -2 Log Likelihood: 431.219								
d. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.								

جدول شماره هشتم، مدل را با مدل صفر مقایسه می‌کند. اگر sig معنی‌دار باشند مدل‌های پیش‌بینی توانسته‌اند مدل را بهبود بخشند.

جدول شماره هشتم

Omnibus Tests of Model Coefficients				
		Chi-square	df	Sig.
Step 1	Step	23.495	5	.000
	Block	23.495	5	.000
	Model	23.495	5	.000

در جدول شماره نهم میزان R Squar مهم است و معمولا Nagelkerke R Square گزارش می شود و بین ۰-۱ متغیر است ولی Cox & Snell R Square تفسیرش مشکل است و هیچ وقت به یک نمی رسد. بدین ترتیب ۰/۰۹۷ مدل توسط متغیرهای مستقل وارد شده تبیین می شود.

جدول شماره نهم

Model Summary			
Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	407.724 ^a	.072	.097

a. Estimation terminated at iteration number 4 because parameter estimates changed by less than .001.

جدول شماره دهم کی دو را نشان می دهد که بهتر است غیرمعنی دار باشد زیرا اگر معنی دار باشد مدل و داده ها مشکل دارند در اینجا مدل غیر معنی دار می باشد و مدل ما درست است.

جدول شماره دهم

Hosmer and Lemeshow Test			
Step	Chi-square	df	Sig.
1	4.205	8	.838

جدول شماره یازدهم یک جدول برای تخمین بالاست.

جدول شماره یازدهم

Contingency Table for Hosmer and Lemeshow Test						
		TALAG = Non-probability		TALAG = Probability		Total
		Observed	Expected	Observed	Expected	
Step 1	1	24	22.935	7	8.065	31
	2	21	22.821	19	17.179	40
	3	18	17.260	14	14.740	32
	4	17	14.848	12	14.152	29
	5	11	13.513	20	17.487	31
	6	11	12.476	21	19.524	32
	7	6	6.488	11	10.512	17
	8	18	17.165	36	36.835	54
	9	6	7.202	17	15.798	23
	10	10	7.293	14	16.707	24

همانطور که در جدول شماره دوازدهم می‌بینیم میزان overall percentage بهبود یافته. هرچه فاصله بیشتر باشد مدل قدرت پیش‌بینی بالایی دارد، که در مدل ما فاصله تقریباً زیاد است در مدل صفر ۵۴ و در اینجا ۶۳ شده است.

جدول شماره دوازدهم

Classification Table ^a					
	Observed	Predicted			Percentage Correct
		TALAG			
		Non-probability	Probability		
Step 1	TALAG	Non-probability	79	63	55.6
		Probability	52	119	69.6
	Overall Percentage				63.3

a. The cut value is .500

جدول شماره سیزدهم جدول مهمی می‌باشد چراکه تاثیر متغیرهای مستقل را نشان می‌دهد. اول به معناداری متغیرها نگاه می‌کنیم اگر معنادار بود سراغ نسبت برتری ($Exp(b)$) می‌رویم اگر بیش از یک باشد احتمال وقوع زیاد و اگر کمتر از یک باشد احتمال وقوع کمتر می‌باشد که در این مدل تحصیلات و وضعیت فعالیت معنی‌دار

آمار پیشرفته در علوم اجتماعی

می‌باشند. تحصیلات ۲/۹۸ احتمال نگرش مثبت به طلاق را افزایش می‌دهد. وضعیت فعالیت نیز ۰/۶۷ احتمال نگرش مثبت به طلاق را افزایش می‌دهد.

جدول شماره سیزدهم

Variables in the Equation									
		B	S.E.	Wald	df	Sig.	Exp(B)	95.0% C.I. for EXP(B)	
								Lower	Upper
Step 1 ^a	vazeatfaaleat	-.560	.295	3.606	1	.058	.571	.320	1.018
	jens	.316	.271	1.355	1	.244	1.371	.806	2.334
	tagiremahlezendegi	-.022	.274	.006	1	.936	.978	.572	1.673
	tahsilat	1.094	.318	11.828	1	.001	2.986	1.601	5.569
	vaziytetaahol	.057	.234	.060	1	.807	1.059	.670	1.674
	Constant	-2.069	1.041	3.950	1	.047	.126		

a. Variable(s) entered on step 1: vazeatfaaleat, jens, tagiremahlezendegi, tahsilat , vaziytetaahol.

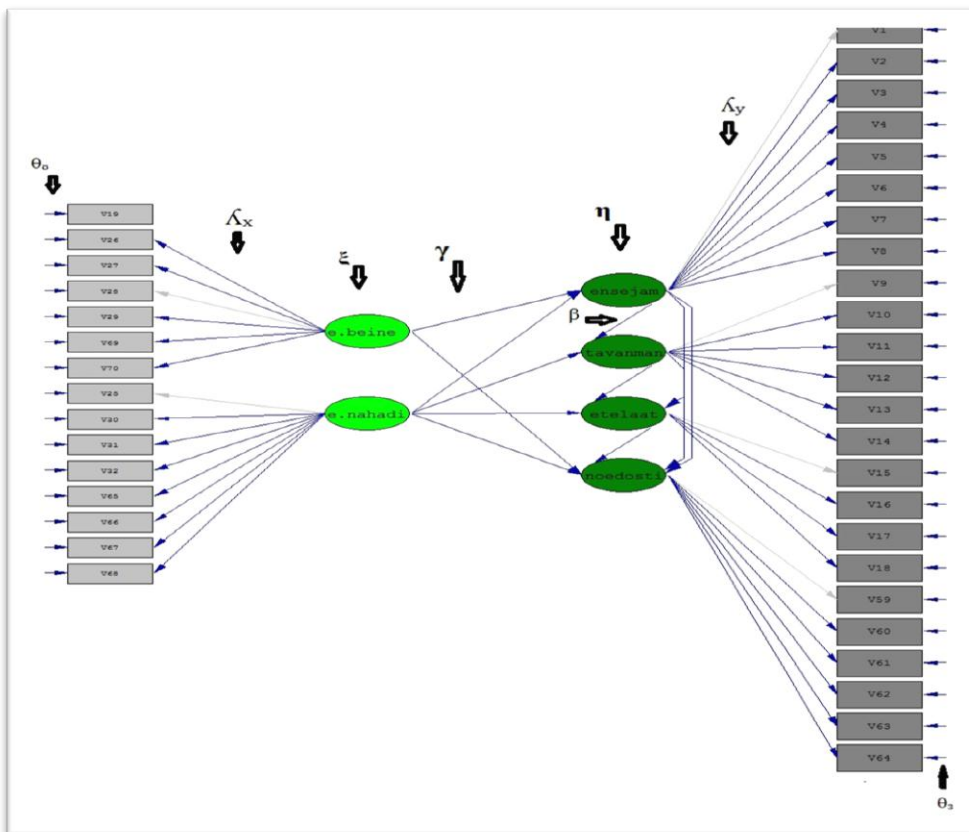
جدول چهاردهم همبستگی متغیرهای مستقل را نشان می‌دهد. متغیرهای مستقل بهتر است با یکدیگر همبستگی نداشته باشند.

جدول شماره چهاردهم

Correlation Matrix							
		Constant	vazeatfaaleat	jens	tagiremahlezendegi	tahsilat	vaziytetaahol
Step 1	Constant	1.000	-.007	-.413	-.545	-.734	-.494
	vazeatfaaleat	-.007	1.000	.427	.000	-.228	-.421
	jens	-.413	.427	1.000	.071	.005	-.232
	tagiremahlezendegi	-.545	.000	.071	1.000	.094	.016
	tahsilat	-.734	-.228	.005	.094	1.000	.515
	vaziytetaahol	-.494	-.421	-.232	.016	.515	1.000

انواع پارامترها در یک مدل کامل لیزرل:

۱. بتا (β): روابط علی بین $ETA(\eta)$ را اندازه می‌گیرد.
۲. گاما (γ): تاثیر $kxi(\xi)$ بر ETA را نشان می‌دهد.
۳. لامبدا (λ): لامبدا x ، بار عاملی kxi و لامبدا y بارهای عاملی Eta را نشان می‌دهند.
۵. زتا (ζ): خطای اندازه‌گیری هر Eta می‌باشد.
۶. فای (ϕ): همبستگی و هم‌تغییری بین kxi را نشان می‌دهد.
۷. سای (ψ): یکسری روابط همبستگی بین زتاها را می‌گویند.
۸. تتا دلتا (θ_8): خطای اندازه‌گیری مربوط به هر یک از x ها.
۹. تتا اپسیلون (θ_3): خطای اندازه‌گیری مربوط به هر یک از y ها.



پارامترهای آزاد، ثابت و محدود شده

پارامترهای آزاد، پارامترهایی اند که محقق قصد برآورد آنها را دارند و پارامترهای مجهولند. پارامترهای ثابت، معمولا دارای مقدار مشخص و تعیین شده اند و برخی اوقات لازم می شود پارامترهای خاصی را در یک مقدار مشخصی ثابت فرض کنیم. عمدتا ضریب آن یک است. پارامترهای محدود، پارامترهای مجهولی اند که علاقه ای به برآورد آنها وجود ندارند، برحسب پیش فرض ها امکان برآورد آنها وجود ندارد.

انواع ماتریس در مدل لیزرل

Ful matrix: کل اجزای ماتریس آزاد هستند معمولا شکل ماتریس های $\Lambda_x, \Lambda_y, \beta, \gamma$ به این صورت است.

Dignon matrix: یا همان ماتریس قطری که فقط قطر ماتری آزاد هستند. معمولا شکل ماتریس های تتا اسیلون و تتا دلتا به این صورت است.

Symmetric matrix: قطر ماتریس و یک طرف (بالا یا پایین) ماتریس آزاد است. معمولا شکل ماتریس پارامترهای فای و سای به این صورت است.

مدل کامل لیزرل توسط سه معادله تعریف می شود:

$$\eta = \beta\eta + \gamma\xi + \xi \quad \text{مدل تابع ساختاری:}$$

$$y = \lambda_y \eta + \beta\gamma \quad \text{مدل اندازه گیری متغیرهای Y:}$$

$$x = \lambda_x \xi + \delta x \quad \text{مدل اندازه گیری متغیرهای X:}$$

پیش فرض های معادلات ساختاری کواریانس:

۱. روابط خطی بین متغیرها

۲. نرمال بودن توزیع

۳. مشاهدات مستقل

۴. عدم همبستگی بین زتای و کسای

۵. معکوس ماتریس بتا نامنفرد است.

مراحل ساخت مدل ساختاری:

۱. بسط مدل ساختاری و اندازه گیری براساس تئوری و تعیین روابط بیم متغیرها
۲. ساخت مدل مسیر و روابط علی از تئوری‌ها و با استفاده از مفاهیم
۳. انتخاب نوع ماتریس داده و برآورد مدل
۴. ارزیابی همانندی مدل ساختاری: آیا مدل در نظر گرفته شده به سطحی می‌رسد که برآزش‌های با ثباتی داشته باشد؟

عواملی که باعث بی‌ثباتی (**not converge**) شدن مدل می‌شود:

۱. وقتی یک یا چند مورد خطاهای استاندارد بزرگ باشند.
۲. برآوردهای غیر منطقی: منظور ضرایبی که بدست می‌آید دور از عرف باشد مثلا ضرایب با همبستگی خیلی بالا.
۳. پایین بودن درجه آزادی: استفاده از تاثیرات دو سویه متعدد
۴. تعداد زیاد اجزای ضرایب برآوردی در مقایسه با تعداد اجزای ماتریس کواریانس یا همبستگی.

نحوه اجراء مدل لیزرل:

برای اجرای یک مدبل لیزرل بهتر است داده‌ها را در spss ویرایش کنید و داده‌های آماده را در لیزرل import کنید. از منوی statistics ، output options را انتخاب می‌کنید و یکسری اطلاعات کلی در مورد داده‌ها به شما می‌دهد. در مرحله بعد از منوی new .file را انتخاب می‌کنید و در اینجا انتخاب می‌کنید که به چه روشی می‌خواهید مدل را اجراء کنید: path digram , lisrel project ، syntax only . جدول زیر نمونه از syntax مدل لیزرل است.

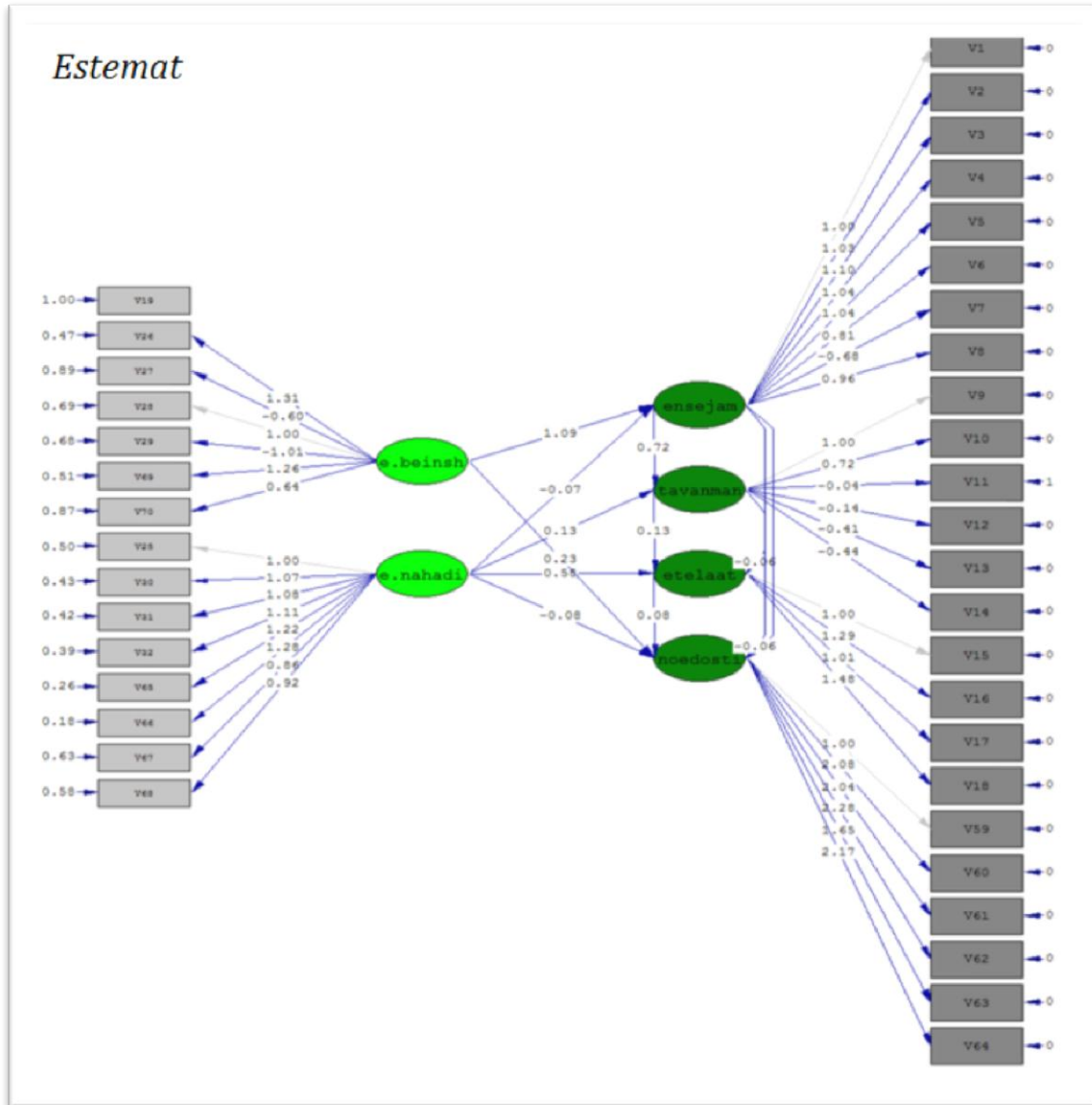
```

TI raftarshahrvandi
!DA NI=39 NO=161 NG=1 MA=KM
SY='E:\LISREL\myLisrel.dsf' NG=1
SE
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18
34 35 36 37 38 39 22 23 24 25 26 27 28 29 30 31 32 33 19 20 21 /
MO NX=15 NY=24 NK=2 NE=4 LY=FU,FI LX=FU,FI BE=FU,FI GA=FU,FI PH=SY,FR PS=DI,FR TE=DI,FR
TD=DI,FR
LE
ensejam tavanmand etlaat noedosti
LK
e.binsh e.nahadi
FR LY(2,1) LY(3,1) LY(4,1) LY(5,1) LY(6,1) LY(7,1) LY(8,1) LY(10,2) LY(11,2)
FR LY(12,2) LY(13,2) LY(14,2) LY(16,3) LY(17,3) LY(18,3) LY(20,4) LY(21,4) LY(22,4)
FR LY(23,4) LY(24,4) LX(2,1) LX(3,1) LX(4,1) LX(6,2) LX(7,2) LX(8,2) LX(9,2)
FR LX(10,2) LX(11,2) LX(12,2) LX(14,1) LX(15,1) BE(2,1) BE(3,1) BE(3,2) BE(4,1)
FR BE(4,2) BE(4,3) GA(1,1) GA(1,2) GA(2,2) GA(3,2) GA(4,1) GA(4,2)
VA 1.00 LY(1,1) LY(9,2) LY(15,3) LY(19,4) LX(1,1) LX(5,2)
PD
OU EF TV SS TO MR MI RS SE

```

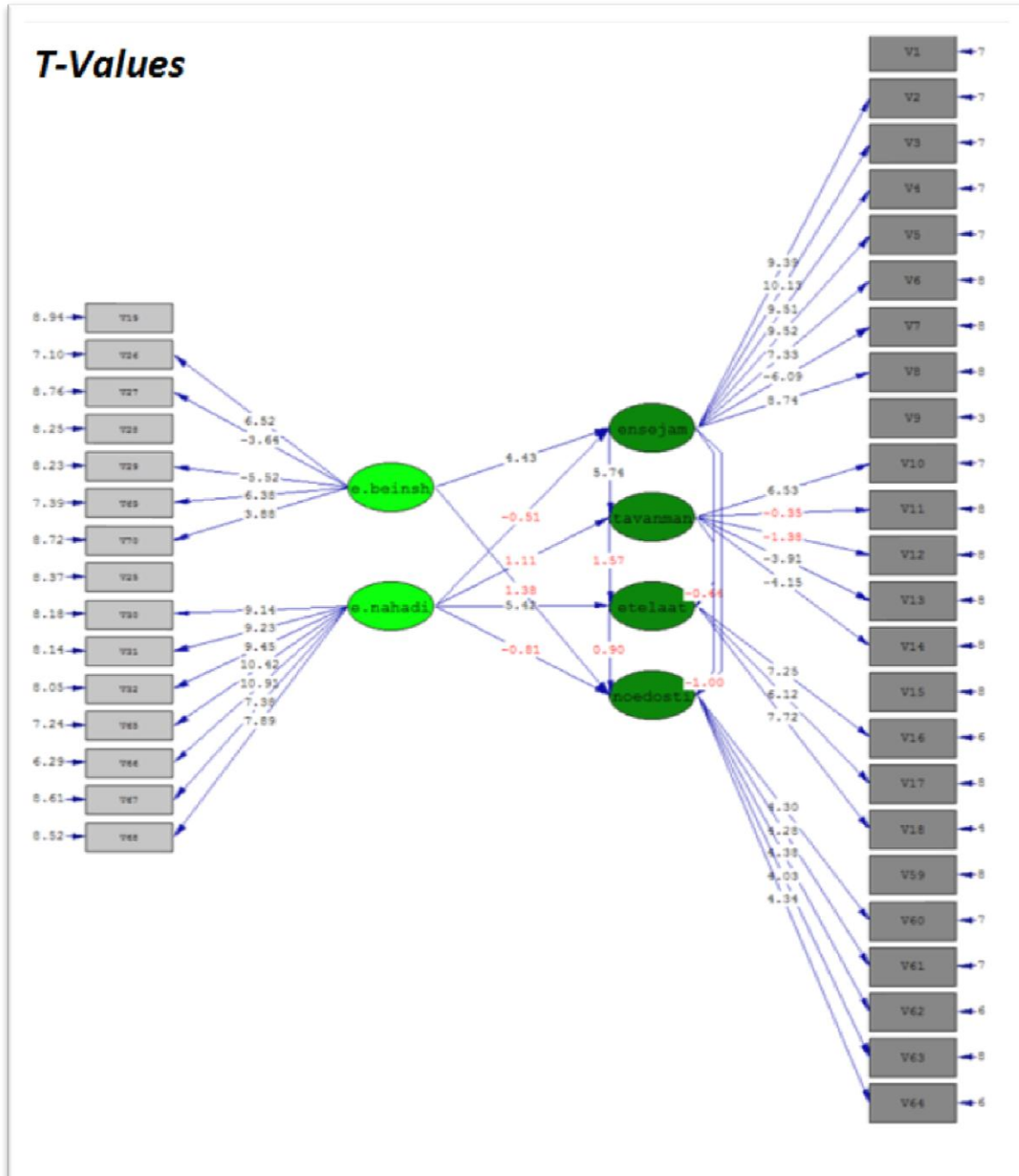
تفسیر مدل لیزرل:

پس از اجراء لیزرل نتایج به صورت دیاگرام (project pth) و همچنین بصورت جدول در فایل poroject out ظاهر می شود. فایل project out در چند جدول اطلاعات کلی و یکسری برآوردهای مقدماتی را نشان می دهد و سپس سراغ برآزش های نهایی می رود. بارهای عاملی هر یک از گویه ها، ضرایب مسیر متغیرهای نهفته، شاخص های برآزش مدل، مقدار باقی مانده ها و در نهایت کمیت های t از جمله جداول مهمی است که معمولا در گزارش ها می آید. نمودار زیر نمودار Estemat بارهای هر یک از گویه ها و ضرایب مسیر متغیرهای نهفته را مشخص می کند. همانطور که مشاهده می شود برخی از عامل ها داری بار عاملی بالاتر از یک می باشد این بدین دلیل است که گویه ای که ما به عنوان گویه ثابت انتخاب کردیم مهمترین گویه نبوده و تاثیر برخی از گویه ها که بار عاملی بالای یک دارند بیشتر است.



با توجه به اینکه در هر متغیر یکی از گویه‌ها را با عدد یک ثابت کرده بودیم تا پایه‌ای برای اندازه‌گیری باشد در نمودار Estemat بارهای عاملی گویه‌ها برحسب آن گویه محاسبه شده‌اند نمودار Standardized Solution بارهای عاملی گویه‌ها و ضرایب مسیر متغیرهای نهفته را بصورت استاندارد شده نشان می‌دهد. بارهای عاملی بالای ۰/۴ نشان می‌دهد گویه‌ها بخوبی توانسته‌اند متغیرها را تبیین کنند همچنین بارهای منفی نشان می‌دهد که گویه‌ها هم‌جهت نیستند. همانطور که در تصویر می‌بینید گویه‌های مربوط به توانمندی‌سازی (Eta2) به جز دو گویه اول بقیه گویه‌ها کمتر از ۰/۴ و منفی می‌باشند. گویه اول نوع دوستی نیز کمتر از ۰/۴ می‌باشد در Ksiها نیز گویه‌های ۲ و ۶ مربوط به اعتماد بین شخصی بارعاملی کمی دارند. در تحلیل ضرایب مسیر(تاثیرات گاما و

T-Value معنی داری همبستگی‌ها و تاثیرات مورد نظر ما را نشان میدهد که باید همواره مقدار آن بالاتر از ۱/۹۶ باشد در اینجا نیز مقادیر پایین تر از ۱/۹۶ و همچنین موارد منفی غیرمعنی داری رابطه را به ما نشان می دهد.



اعدادی که با رنگ قرمز مشخص شده‌اند نشان از غیرمعنی داری مقادیر t دارند. در مدل اندازه‌گیری همه روابط به جز تاثیر توانمندسازی بر روی گویه‌های ۱۱ و ۱۲، معنادار می‌باشند. در مدل ساختاری، رابطه بین اتاها (β)

به جز تاثیر انسجام بر توانمند سازی همه تاثیرات غیر معنی دارند. همچنین در تاثیر کسایها بر اتاها (Y) به جز تاثیر اعتماد بین شخصی بر انسجام و تاثیر اعتماد نهادی بر اطلاعات تمامی روابط غیر معنی دارند و تاثیر معنی-داری ندارند.

نیکویی برآزش مدل

قسمت دوم ارزیابی مدل، شامل برآورد نیکویی برآزش کلی مدل با داده‌های مشاهده شد توسط معیارهای نیکویی برآزش است. معمولا معیار برآزش مدل کمیت کی دو می‌باشد که مدل موجود را نسبت به یک تطابق داده‌ها و مدل مفهومی ما ارزیابی و آزمون می‌کند. اگر مدل مناسب باشد کی دو غیر معنی‌دار می‌آید و این بدین معنی است که مدل و داده‌ها باهم تطابق دارند. در اینجا با توجه به اطلاعات **project out**، کی دو غیر معنی-دار است.

با توجه به این که کی دو شدیداً به حجم نمونه حساس است وقتی نمونه بزرگتر باشد تفاوت‌های جزئی را هم معنی‌دار نشان می‌دهد معمولا از معیارهای دیگری استفاده می‌کنند. معیارهای برآزش مدل به دو دسته کلی تقسیم می‌شوند:

۱. یکسری معیارها مدل شما را با معیارهای مفهومی می‌سنجند.

۲. یکسری معیارها مدل شما را با مدل صفر مقایسه می‌کنند.

مهمترین شاخص‌های برآزش مدل که می‌توان در ارزیابی یک مدل لیزرل در گزارش‌ها آورد:

۱. **Goodness of fit index**: عددی بین صفر و یک می‌باشد هر چه به یک نزدیک‌تر باشد نشان از مطلوبیت مدل دارد و اگر زیر ۰/۹ باشد نامطلوبیت مدل را نشان می‌دهد. در مدل ما این عدد ۰/۶۲ که می‌توان گفت مدل ما مدلی نامطلوب می‌باشد.

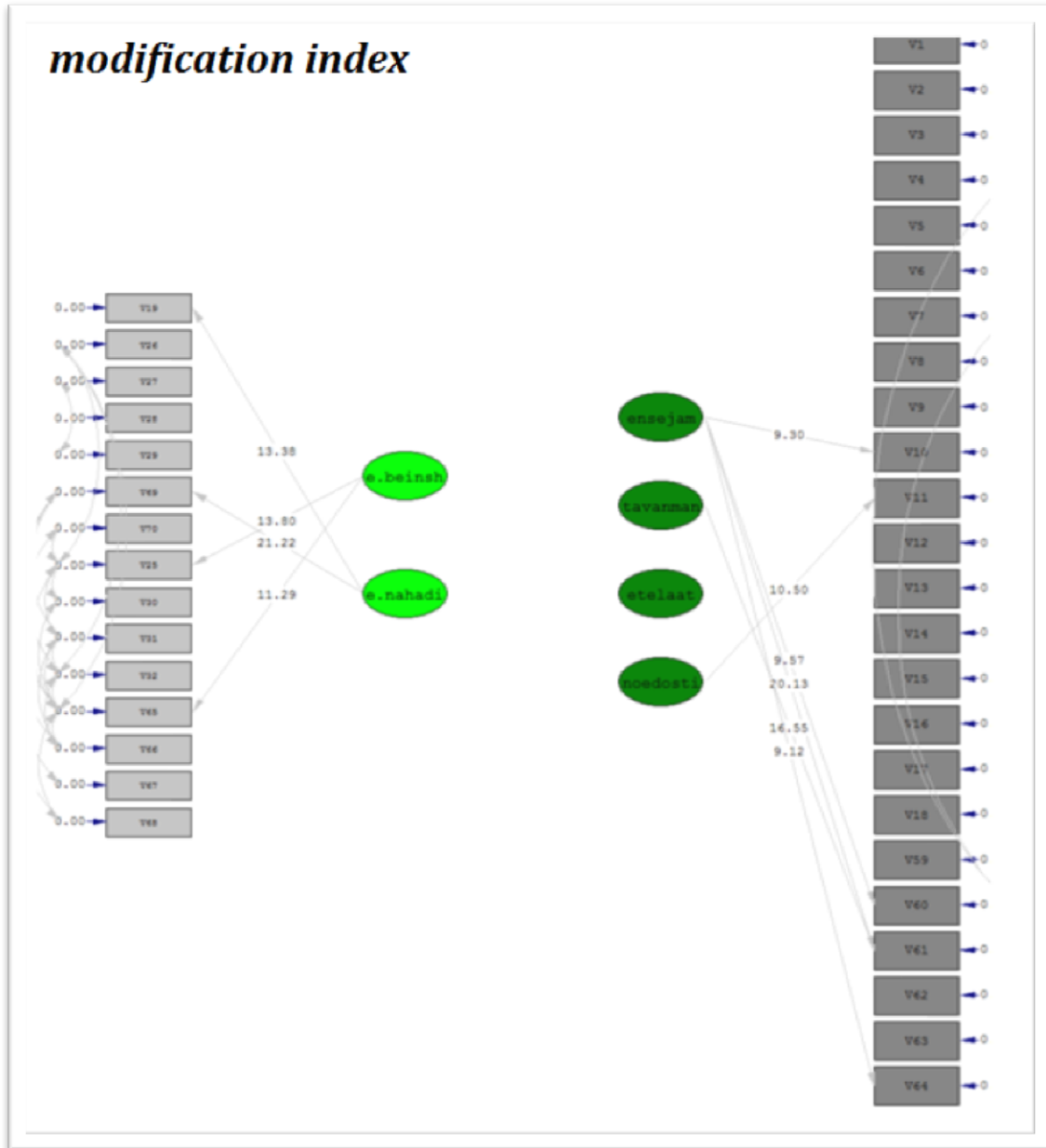
۲. **Ajusted Goodness of fit index**: این شاخص نسبت به درجه آزادی کی دو را تعدیل می‌کند. و همانند GFI باید نزدیک به یک باشد. که در مدل ما ۰/۵۷ است که باز کم است و بیانگر عدم تطابق مدل و داده‌ها.

۳. Root mean square residual: این شاخص هرچه به صفر نزدیک باشد مطلوبیت را می‌رساند و زیر ۰.۰۵ هم نشانگر مناسب بودن مدل است. که در اینجا RMS ما ۰/۰۹ می‌باشد که باز نامطلوبیت مدل را می‌رساند.

۴. همچنین دو شاخص Tucker Lewis Index و Normed Fit Index مطلوبیت مدل را با یک مدل صفر مقایسه می‌کنند که در هر دوی این شاخص‌ها مقادیر بالاتر از ۰/۹ برآزش قابل قبول است. که در اینجا NFI:0/80 و TLI: 0/85 می‌باشد که نشان از نامطلوبیت مدل دارند.

همانطور که گفته شد تمامی شاخص‌ها نشان از نامطلوبیت مدل دارند بنابراین باید مدل را اصلاح کرد. در اینجا توجه به مقادیر Modification Index مهم می‌باشد که در تصویر زیر دیاگرام Modification Index داده‌های لیزرل قابل مشاهده است. رابطه‌یی که MI بالایی داشته باشد شاخص‌های برآزش نیکویی مدل را بهبود می‌بخشد و بهتر است اگر پشتوانه تئوریک دارند این رابطه‌ها آزاد شوند. در مدل حاضر بهتر است رابطه بین اعتماد نهادی و گویه ۶۹ آزاد (Free) شود. باید رابطه‌ها تک تک آزاد شوند.

همچنین توجه به مقادیر باقی‌مانده (Residuals) مهم است، باقی‌مانده‌های خیلی بزرگ و خیلی کوچک (منفی) باید تعدیل شوند. علامت منفی جایی است که Fix کردن یک رابطه موجب بهبود مدل می‌شود. در باقی‌مانده‌های بالاتر از ۲/۵۸ (به عنوان باقی‌مانده‌های خیلی بزرگ) با آزاد کردن رابطه (Free) مدل بهبود می‌یابد. برای تفسیر باقی مانده‌ها و تعدیل مدل بهتر است باقی‌مانده‌های استاندارد (Standardized Residuals) تک تک گویه بررسی شود.





در کانال تلگرام کارنیل هر روز انگیزه خود را شارژ کنید 😊

<https://telegram.me/karnil>

