

آموزش

همبستگی تفکیکی

با نرم افزار SPSS

(Partial correlation)

تدوین: رامین کریمی

www.Kharazmi-Statistics.ir

اصولا وقتی همبستگی بین دو متغیر محاسبه می شود، تفسیر آن بحث برانگیز و ابهام آمیز است. این پرسش مطرح می شود که همبستگی بدست آمده چه چیز را توجیه می کند؟ آیا رابطه بین متغیرها مستقیم و یا ناشی از نفوذ متغیر دیگری است؟ تحلیل همبستگی جزئی تلاش می کند برخی از ابهامات موجود در تفسیر همبستگی را حل کند. به منظور آن که بتوانیم رابطه بین زوج معینی از متغیرها را ارزشیابی کنیم، گاه ضرورت دارد اثر متغیر سوم را حذف کنیم. برای مثال، انتظار می رود بین گنجینه واژگان و اندازه کفش، در میان کودکان خردسال همبستگی مثبت قابل ملاحظه ای وجود داشته باشد. اما تردید وجود دارد که بین این دو متغیر واقعا رابطه علت و معلولی وجود داشته باشد، زیرا پشت این دو متغیر، متغیر سوم یعنی سن وجود دارد. با افزایش سن، هر دو متغیر اندازه کفش و خزانه واژگان افزایش می یابد. بنابراین برای تعیین رابطه واقعی بین دو متغیر، سن باید کنترل شود. متغیری که حذف شده یا ثابت نگه داشته می شود، متغیر کنترل نامیده می شود. اگر بتوانیم نفوذ سن را حذف کنیم. انتظار می رود همبستگی بین این دو متغیر عملا صفر باشد. با انتخاب آزمودنی های هم سن و سال می توان این عامل را به سادگی کنترل کرد. در این صورت واریانس متغیر سن به صفر یا نزدیک به صفر کاهش می یابد. روش محاسبه همبستگی جزئی (تفکیکی)، یکی دیگر از روش های مهم کنترل (آماری) است. ضریب همبستگی جزئی بیان گر رابطه بین دو متغیر در حالتی است که نفوذ یک یا چند متغیر دیگر حذف شده باشد.

ضریب همبستگی برای داده های فاصله ای نشان دهنده ی شدت رابطه ی دو متغیر است و این ضریب به ما می گوید اگر نمره افراد را در یک متغیر بدانیم، با چه دقتی می توانیم نمره آن ها را در متغیر دیگر پیش بینی کنیم. ضریب همبستگی جزئی هم نشان دهنده همین اطلاعات است، منتهی با از بین بردن اثر متغیرهای مستقل دیگری که حذف شده اند. با محاسبه ضریب همبستگی جزئی بین هر متغیر مستقل و متغیر وابسته و مقایسه آن ها معلوم می شود که کدام متغیر مستقل، دقیق ترین پیش بین متغیر وابسته است (یعنی کدام متغیر مستقل دارای قوی ترین رابطه با متغیر وابسته است). ضریب همبستگی جزئی عمدتا در مورد متغیرهای فاصله ای به کار می رود و مبتنی بر ضریب همبستگی پیرسون است.

پیش فرض ها

- تمامی متغیرها باید دارای مقیاس کمی (فاصله ای / نسبی) باشند. یعنی هم دو متغیر اصلی و هم متغیرهای کنترل باید کمی باشند.

- متغیرهای مورد آزمون باید رابطه خطی با همدیگر داشته باشند.

نکته: همبستگی جزئی (تفکیکی) تنها برای مدل های کوچک مفید است. یعنی مدل هایی که ۳ یا ۴ متغیر را در بر می گیرند. بنابراین، برای مدل های بزرگ، زمانی که داده ها فاصله ای / نسبی و یا نزدیک به آن ها بودند، می توانیم از تحلیل مسیر یا مدل سازی معادلات ساختاری و زمانی که مقیاس متغیرها در سطح سنجش کیفی (اسمی / ترتیبی) بودند، از مدل سازی لگاریتم خطی استفاده می کنیم.

مفاهیم کلیدی

متغیر آزمون: به متغیر که کنترل می شود، می گویند. در واقع، همان متغیر کنترل است.

رابطه مرتبه صفر: به رابطه اولیه بین دو متغیر قبل از کنترل آماری اطلاق می شود که مقدار آن را ضریب همبستگی ساده (پیرسون) برابر است.

رابطه مشروط (تفکیکی/جزئی): به رابطه بین دو متغیر بعد از حذف اثر متغیر کنترل گفته می شود. این رابطه مشروط می تواند به صورت مرتبه های اول، دوم و ... باشد:

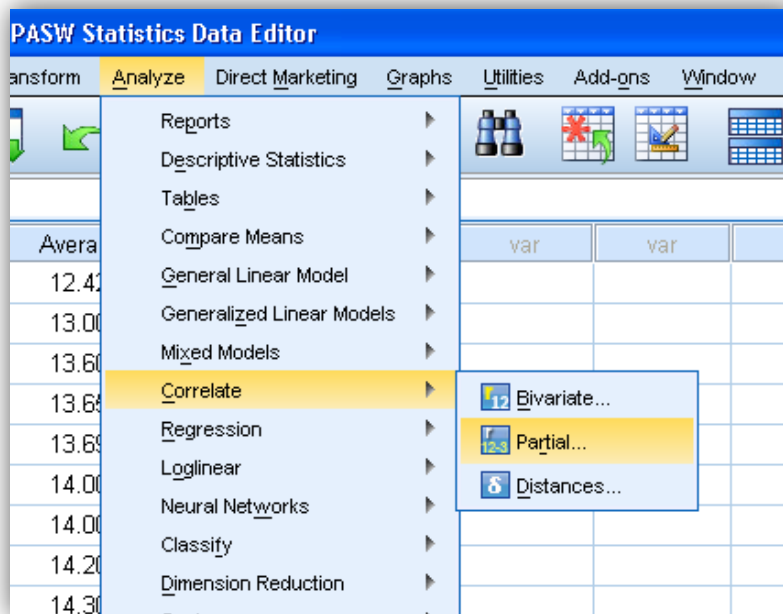
۱- رابطه مشروط مرتبه اول: به رابطه بین دو متغیر بعد از حذف اثر فقط یک متغیر کنترل گفته می شد.

۲- رابطه مشروط مرتبه دوم: به رابطه بین دو متغیر بعد از حذف اثر همزمان دو متغیر کنترل گفته می شود.

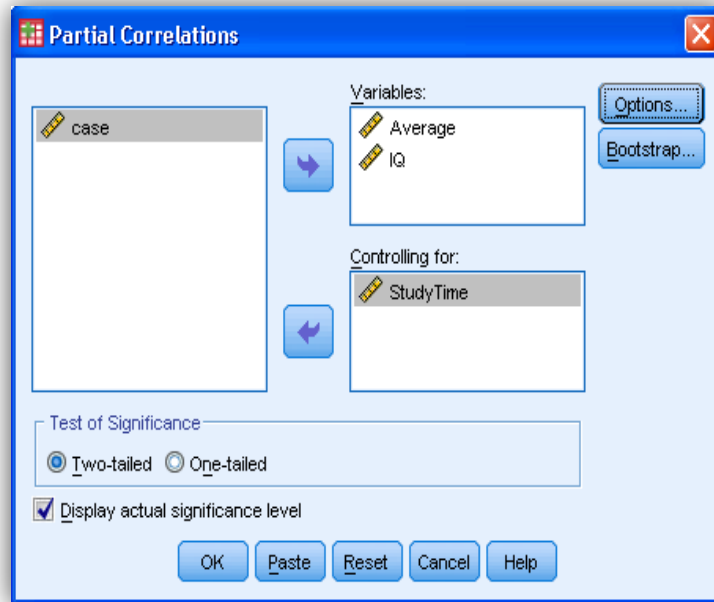
مثال: می خواهیم بدانیم که آیا رابطه بین معدل تحصیلی و ضریب هوشی دانش آموزان واقعی است و یا وجود بین این دو متغیر با کنترل متغیر ساعات مطالعه دانش آموزان از بین می رود. به عبارتی می خواهیم پی ببریم که آیا با کنترل متغیر سومی به نام تعداد ساعات مطالعه دانش آموزان، رابطه بین متغیرهای معدل و ضریب هوشی آنان همچنان معنادار خواهد بود یا خیر؟ (آیا علی رغم فرض ما، تعداد ساعات مطالعه دانش آموزان است که نمره آنان را تعیین می کند؟)

اجرا

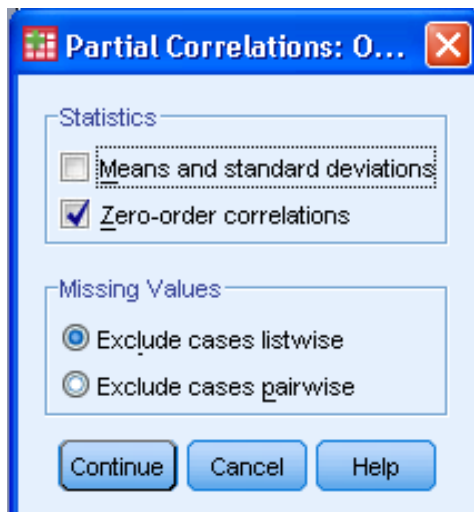
۱- مسیر Analyze > Correlate > Partial... را اجرا می کنیم.



۲- در پنجره ای که باز می شود، متغیرهای معدل تحصیلی و ضریب هوشی (Average و IQ) را وارد کادر Variables می کنیم و متغیر ساعات مطالعه (Study Time) را که نقش متغیر کنترل را دارد، وارد کادر Controlling for می نمایم.



۳- در کادر اصلی بر روی Options کلیک کرده و گزینه Zero-order correlation را انتخاب می کنیم. در انتها Continue و در نهایت OK را می زنیم.



Control Variables		Average	IQ	StudyTime	
none ^a	Average	Correlation	1.000	.963	.970
		Significance (2-tailed)	.	.000	.000
		df	0	28	28
	IQ	Correlation	.963	1.000	.997
		Significance (2-tailed)	.000	.	.000
		df	28	0	28
StudyTime	Correlation	.970	.997	1.000	
	Significance (2-tailed)	.000	.000	.	
	df	28	28	0	
StudyTime	Average	Correlation	1.000	-.180	
		Significance (2-tailed)	.	.351	
		df	0	27	
	IQ	Correlation	-.180	1.000	
		Significance (2-tailed)	.351	.	
		df	27	0	

a. Cells contain zero-order (Pearson) correlations.

همبستگی مرتبه صفر
(بدون کنترل متغیر دیگر)

همبستگی مرتبه اول
(با کنترل متغیر سوم یعنی
تعداد ساعات مطالعه)

۱- اگر رابطه جداگانه بین دو متغیر معنی دار شد، در عین حال رابطه آنها با وجود متغیر کنترل، باز هم معنی دار شد، در آن صورت رابطه واقعی است (در چنین حالتی، احتمال وجود رابطه بین متغیر کنترل با هر دو متغیر اصلی کم است).

۲- اگر رابطه جداگانه بین دو متغیر معنی دار شد، اما رابطه آنها با وجود متغیر کنترل معنی دار نشد، در آن صورت رابطه کاذب است. (در چنین حالتی احتمال وجود رابطه بین متغیر کنترل با هر دو متغیر اصلی زیاد است). بنابراین می توان چنین تصور کرد که رابطه مشاهده شده بین دو متغیر واقعی نبوده و وجود چنین رابطه ای به خاطر وجود رابطه ای است که بین هر کدام از این دو متغیر اصلی و متغیر کنترل برقرار است.

نکته: چنانچه رابطه جداگانه بین دو متغیر معنی دار شد، در عین حال رابطه آنها با وجود متغیر کنترل، باز هم معنی دار شد، اما شدت همبستگی بین دو متغیر بعد از ورود متغیر کنترل کمتر شد (همبستگی مرتبه صفر بیشتر از همبستگی مرتبه اول، اما هر دو معنادار)، این بدان معناست که بخشی از همبستگی اولیه ناشی از همبستگی هر دو متغیر با متغیر سوم است. به عبارت دیگر، متغیر سوم بخشی از همبستگی اولیه را تبیین یا تفسیر می کند.

تفسیر: در نمونه اجرا شده (جدول بالا Correlations)، همان طور که مشاهده می کنیم، قبل از کنترل متغیر تعداد ساعات مطالعه (none)، رابطه بین دو متغیر ضریب هوشی و معدل افراد معنادار است ($P < .01$) و شدت رابطه بین دو متغیر نیز برابر است با $.963$. که به این معناست که بین دو متغیر رابطه معناداری وجود دارد. اما بعد از کنترل متغیر سوم (ساعات مطالعه)، مشاهده می کنیم که رابطه مرتبه صفر دیگر معنادار نیست ($P = .351$). به عبارتی با کنترل متغیر ساعات مطالعه، رابطه اولیه بین دو متغیر ضریب هوشی و معدل از بین رفت و این گونه تفسیر می شود که رابطه اولیه کاذب است.

منابع

- حبیب پور، کرم و صفری، رضا (۱۳۸۸) راهنمای جامع کاربرد SPSS در تحقیقات پیمایشی، تهران: نشر لویه.
- دواس، دی. ای (۱۳۷۶)، پیمایش در تحقیقات اجتماعی، ترجمه هوشنگ نایی، تهران: نشر نی.
- نایی، هوشنگ (۱۳۸۷)، برنامه کامپیوتری آمار در علوم اجتماعی SPSS، تهران: انتشارات سروش.
- هومن، حیدر علی (۱۳۹۰)، تحلیل داده های چند متغیری در پژوهش رفتاری، تهران: پیک فرهنگ.

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"