

فصل هشتم

آزمون‌های آماری ناپارامتریک

هدف‌های یادگیری

- از دانشجو انتظار می‌رود که پس از خواندن فصل هشتم بتواند:
۱. تفاوت آزمون‌های آماری پارامتریک و ناپارامتریک را بیان کند.
 ۲. محاسن و معایب آزمون‌های ناپارامتریک را توضیح دهد.
 ۳. آزمون کای دو را محاسبه کند.
 ۴. کاربرد، محاسن و معایب آزمون لمان - ویتنی را بیان کند.
 ۵. آزمون لمان - ویتنی را محاسبه کند.
 ۶. کاربرد، محاسن و معایب آزمون کروسکال - والیس را بیان کند.
 ۷. آزمون کروسکال - والیس را محاسبه کند.
 ۸. کاربرد، محاسن و معایب آزمون ویلکاکسون را بیان کند.
 ۹. آزمون ویلکاکسون را محاسبه کند.
 ۱۰. کاربرد، محاسن و معایب آزمون فریدمن را بیان کند.
 ۱۱. آزمون فریدمن را محاسبه کند.

آزمون‌های آماری ناپارامتریک

در پژوهش غالباً داده‌هایی جمع‌آوری می‌شود که مقیاس اندازه‌گیری آن‌ها اسمی یا رتبه‌ای است یا گاهی متغیری مورد اندازه‌گیری قرار می‌گیرد که مقیاس آن فاصله‌ای است ولی توزیع آن در جامعه نرمال نیست. در چنین شرایطی پژوهشگر ملزم به استفاده از آزمون‌های ناپارامتریک است. استفاده از آزمون‌های ناپارامتریک به هیچ پیش‌فرضی

مشخصی نیاز ندارد ولی انتخاب آزمون‌های پارامتریک به پیش‌فرض‌های مشخصی نیاز دارد و در صورتی که این پیش‌فرض‌ها تحقق نیابد نمی‌توان از آزمون پارامتریک استفاده کرد. مهم‌ترین پیش‌فرض‌های آزمون پارامتریک چنانچه قبلاً ذکر شد عبارتند از:

۱. مقیاس حداقل فاصله‌ای باشد.

۲. توزیع نمره‌ها در هر دو جامعه نرمال باشد.

۳. واریانس‌ها برابر باشد.

اگر واریانس‌ها برابر نباشد اما تعداد نمونه مساوی باشد از آزمون‌های پارامتریک می‌توان استفاده کرد. اما وقتی واریانس‌ها همگن نباشد و تعداد افراد نمونه هم مساوی نباشد. نمی‌توان از آزمون‌های پارامتریک استفاده کرد و حتماً باید از آزمون ناپارامتریک استفاده شود. (شیولسون ترجمه کیامش جلد دوم ۲۲۳)

در حجم‌های کوچک ($n < 10$) نمونه کمتر از ۱۰ نفر استفاده از آزمون‌های ناپارامتریک، منطقی‌تر و بهتر است. (شیولسون ترجمه کیامش، ۲۱۶:۱۳۸۲).

محاسن آزمون‌های آماری ناپارامتریک

۱. گزاره‌های احتمالاتی که در بیشتر آزمون‌های ناپارامتریک به دست می‌آید «احتمالاً قطعی» هستند.

۲. اگر حجم نمونه به کوچکی ($N=6$) باشد، راهی به جزء استفاده از آزمون‌های ناپارامتریک نیست.

۳. تعدادی از آزمون‌های آماری ناپارامتریک وجود دارند که برای تجزیه و تحلیل نمونه‌هایی مناسب هستند که از مشاهده چندین جامعه مختلف حاصل شده است.

این در حالی است که هیچ‌یک از آزمون‌های پارامتریک نمی‌توانند داده‌های حاصل از این نوع را تجزیه و تحلیل کنند، مگر اینکه یک رشته مفروضات غیر واقعی ارائه کنند.

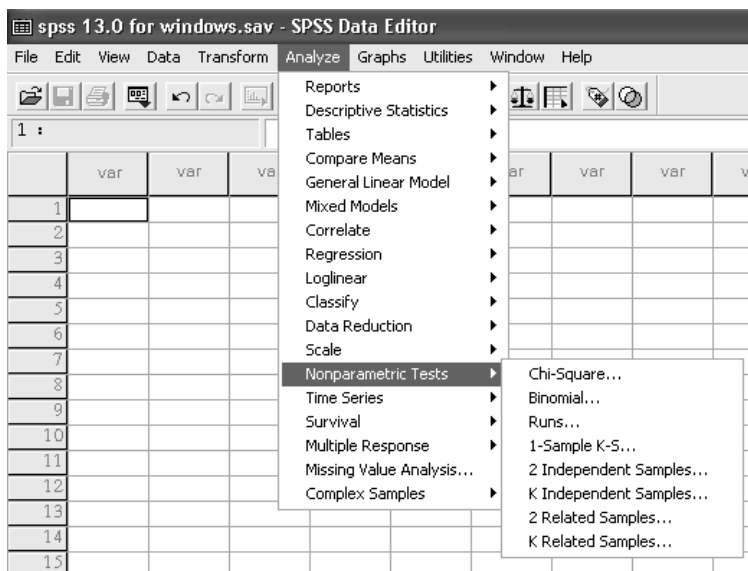
۴. آزمون‌های ناپارامتریک گوناگونی وجود دارند که برای تجزیه و تحلیل داده‌هایی که به صورت رتبه‌ای هستند و همچنین داده‌هایی که نمرات به ظاهر شمارشی آن‌ها رتبه را دارا هستند، مناسبند بدین معنی که محقق ممکن است فقط بتواند بگوید که آزمونی‌های تحقیق او، فلان خصیصه را کم یا زیاد دارند بدون اینکه بتواند بگوید چقدر کمتر یا زیادتر می‌باشد.

۵. آزمون‌های ناپارامتریک از نظر یادگیری و کاربرد به مراتب ساده‌تر و آسان‌تر از آزمون‌های پارامتریک است.

معایب آزمون‌های آماری ناپارامتریک

۱. اگر تمام مفروضات مدل آماری پارامتریک برآورده شوند و اگر مقیاس سنجش ما به حد لزوم یعنی به حد فاصله‌ای برسند، در آن صورت به‌کاربردن آزمون‌های ناپارامتریک اتلاف مقداری از داده‌ها را به همراه دارد. میزان اتلاف به‌وسیله توان - کارآمدی آزمون‌های ناپارامتریک بیان می‌شود.
۲. هنوز روش‌های ناپارامتریک که برای آزمون تعامل‌ها در مدل تحلیل واریانس مناسب باشد ابداع نشده است.
۳. اعتراض دیگر به آزمون‌های ناپارامتریک این است که جدول‌ها مقادیر معنی‌دار مربوط به آزمون‌های ناپارامتریک در منابع مختلف، پراکنده است و اکثر این منابع بسیار تخصصی هستند. بنابراین چنین منابعی به سختی در دسترس علاقه‌مندان قرار می‌گیرد. (سیگل ترجمه کریمی، ۱۳۸۰: ۴۳)

زیر منوی (Nonparametric Test) برای اجرای آزمون‌های ناپارامتری در منوی (Statistics) گنجانده شده است. شکل زیر فرمان‌های مختلف این زیر منو را نشان می‌دهد:



آزمون ناپارامتریک خی دو (Chi-Square)

آزمون χ^2 یکی از آزمون‌های ناپارامتریک مهم است در آزمون‌های ناپارامتریک χ^2 هم برای مقایسه تک نمونه‌ای (نرمال بودن جامعه)، مقایسه دو گروه مستقل و یک متغیر آزمایشی (به جای U مان-وینتی)، برای مقایسه سه گروه یا بیشتر (k گروه) با یک متغیر آزمایشی (به جای کروسکال وایس) به کار می‌رود.

$$\chi^2 = \sum_{i=1}^r \sum_{j=1}^c \frac{(O_{ij} - E_{ij})^2}{E_{ij}}$$

کاربرد

در بعضی تحقیقات، پژوهشگران با مشاهده و شمارش تعداد آزمونی‌ها (یعنی با فراوانی) سرو کار دارند. مثلاً تعداد اعضای احزاب مختلف یا تعداد رأی‌دهندگان زمانی که پژوهشگران با نظر و علاقه و نگرش افراد سر و کار دارد. علاوه بر موارد کاربرد χ^2 که ذکر شد، عمده ترین ملاک که باید به آن توجه شود، مقیاس است. وقتی متغیرها در سطح مقیاس نامی (اسمی) و رتبه‌ای باشند از χ^2 استفاده کنید. (شیولسون ترجمه کیامنش، ۱۳۸۲: ۳۵۴).

معایب

آزمون χ^2 زمانی به کار برده می‌شود که داده‌ها به صورت مقیاس اسمی باشد. هنگامی که مقیاس اندازه‌گیری ترتیبی یا فاصله‌ای است بهتر است از آزمون دیگری استفاده شود. (دلاور، ۱۳۸۰: ۴۹۸)

مثال: فرض کنید در یکی از کشورهای اروپایی، پژوهشگری علاقه‌مند است به چنین پرسشی پاسخ دهد: «آیا در زمینه تهیه بمب اتمی بین افرادی که در ارتش خدمت کرده اند و افرادی که در ارتش خدمت نکرده اند تفاوت عقیده وجود دارد؟» فرض کنید از طریق تصادفی از ۵۰ نفر پرسیده شده که:

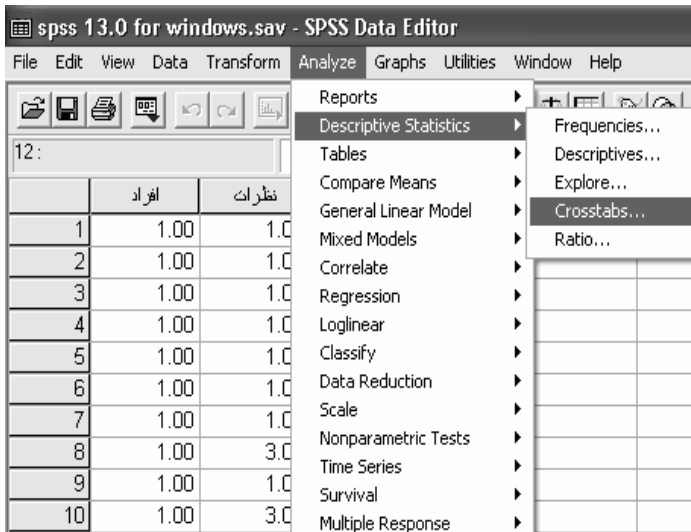
۱. آیا در ارتش خدمت کرده اید؟ بلی، خیر

آزمون‌های آماری ناپارامتریک ۱۳۳

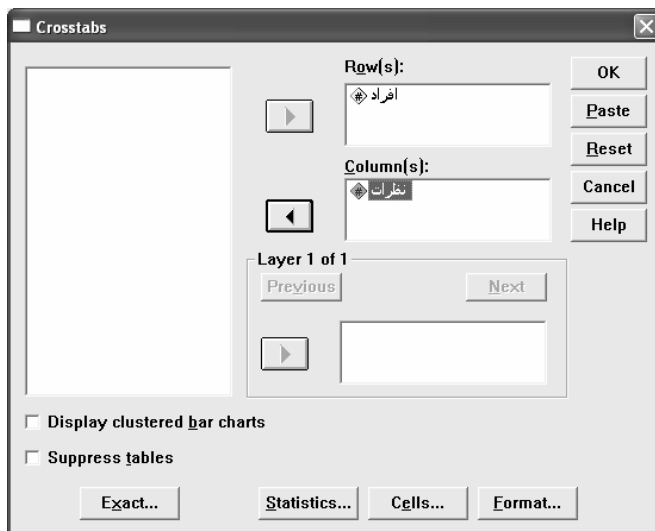
۲. آیا با تهیه بمب اتمی موافق هستید؟ بلی، خیر، بی نظر (شیولسون ترجمه کیامنش، ۳۵۵:۱۳۸۲).

بی نظر	خیر	بلی	نظرات افراد
۵	۶	۱۹	نظامیان
۷	۸	۵	غیرنظامیان

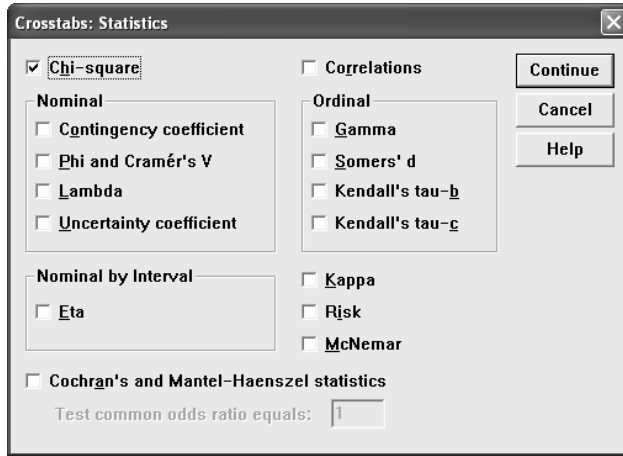
برای اجرای آزمون فوق در Spss مراحل زیر را انجام دهید



بعد از اجرای این فرمان شکل زیر ظاهر می شود:



در پنجره ظاهر شده یکی از متغیرها را به قسمت (Row(s)) و متغیر دیگر را به قسمت (Column(s)) وارد کنید، سپس گزینه (Statistics...) را کلیک کنید تا پنجره زیر باز شود:



در پنجره بالا گزینه (Chi-square) را انتخاب نمایید سپس گزینه (Continue) را فعال نمایید تا به پنجره (Crosstabs) بازگردید در این پنجره گزینه (OK) را کلیک کنید و نتایج به صورت خروجی زیر ظاهر خواهد شد:

Cross tabs

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
افراد * نظرات	50	100.0%	0	.0%	50	100.0%

Crosstabulation افراد * نظرات

		نظرات			Total	
		بلی	خیر	بی نظر		
افراد	در ارتش خدمت کرده	Count	19	6	5	30
		Expected Count	14.4	8.4	7.2	30.0
	در ارتش خدمت نکرده	Count	5	8	7	20
		Expected Count	9.6	5.6	4.8	20.0
Total		Count	24	14	12	50
		Expected Count	24.0	14.0	12.0	50.0

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	7.068 ^a	2	.029
Likelihood Ratio	7.316	2	.026
Linear-by-Linear Association	5.701	1	.017
N of Valid Cases	50		

a.

در جدول (Crosstabulation) تعداد افراد پاسخگو و در جدول (Chi-square Tests) سطح معنی‌داری ذکر شده است که در این مثال، مقدار کای‌دو پیرسون برابر ۷/۰۶ است که این مقدار در سطح ۰/۰۵ معنی‌دار است. بنابراین بین نظرات نظامیان و غیر نظامیان در استفاده از بمب اتمی تفاوت معنی‌داری وجود دارد. و با مقایسه نظرات افراد می‌توان گفت که اکثر افرادی که در ارتش خدمت کرده‌اند، موافق استفاده از بمب اتمی هستند و اکثر افرادی که در ارتش خدمت نکرده‌اند مخالف استفاده از بمب اتمی هستند.

آزمون‌های ناپارامتریک مربوط به دو گروه مستقل (2 Independent samples)

آزمون t ویلکاکسون (من-ویتنی) برای دو گروه مستقل

یکی از آزمون‌های ناپارامتریک مربوط به دو گروه مستقل است که برای مقایسه میانگین دو جامعه، وقتی که مقیاس اندازه‌گیری رتبه‌ای است، مطلوب می‌باشد و در مقابل آزمون پارامتری t بوجود آمده است. زمانی که محقق قصد مقایسه دو گروه مستقل را دارد و فرض نرمال بودن و یکسانی واریانس‌ها برقرار نیست از آزمون U من-ویتنی استفاده می‌شود (شیولسون ترجمه کیامنش، ۱۳۸۲:۲۱۷).

$$z = \frac{2}{\sqrt{\frac{n_1 n_2}{n(n-1)} \left(\frac{n^2 - n}{12} - \sum T \right)}}$$

کاربرد

این آزمون برای تشخیص تفاوت بین دو جامعه (مقایسه میانگین دو جامعه) که با استفاده از نمونه‌های تصادفی از همان جامعه انتخاب شده است و در صورتی که مقیاس

اندازه‌گیری صفت متغیر مورد مطالعه به صورت مقیاس ترتیبی باشد، به‌کار برده می‌شود.

محاسن

آزمون U من- ویتنی در مقایسه با سایر آزمون‌های ناپارامتریک مربوط به دو گروه مستقل قویتر است و می‌توان آن را در آزمون‌های یک دامنه‌ای و دو دامنه‌ای برای آزمایش فرض صفر بکار برد.

معایب

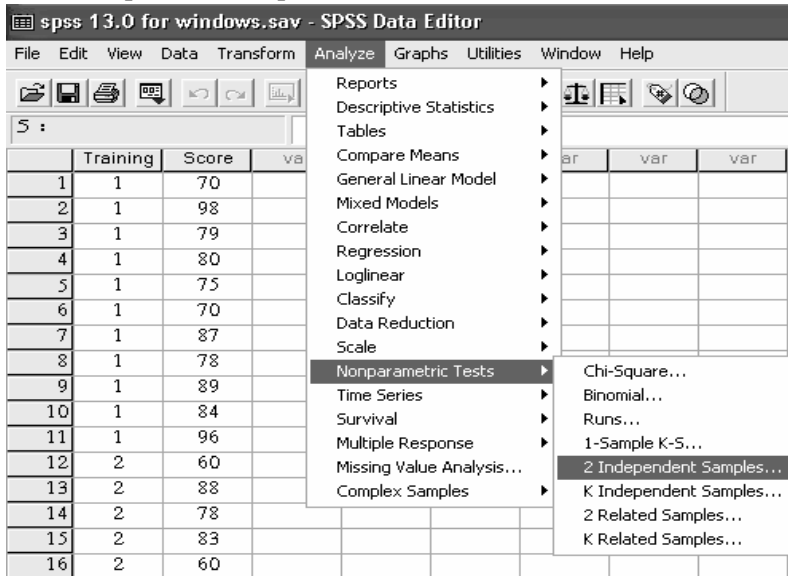
عیب U من- ویتنی این است که تنها می‌توانیم به مقایسه دو میانگین پردازیم و اگر بیش از دو باشد، نمی‌توان از این آزمون استفاده کرد. بلکه باید از آزمون کروسکال والیس استفاده نماییم.

مثال: دو گروه از دانش‌آموزان کلاس درس بهداشت در دوره اول دبیرستان انتخاب شدند و در یک گروه طرز به هوش آوردن مصنوعی بیمار با روش بحث و سخنرانی (Lecture) و گروه دیگر با نمایش فیلم استریپ (Film) آموزش داده شد. در پایان دوره آموزش هر دو گروه با آزمون واحدی مورد آزمایش قرار گرفتند، اکنون با استفاده از آزمون U من- ویتنی در مورد تفاوت دو روش تدریس مورد بحث اظهار نظر کنید (پاشا شریفی نجفی زند، ۱۳۷۵:۲۷۳)

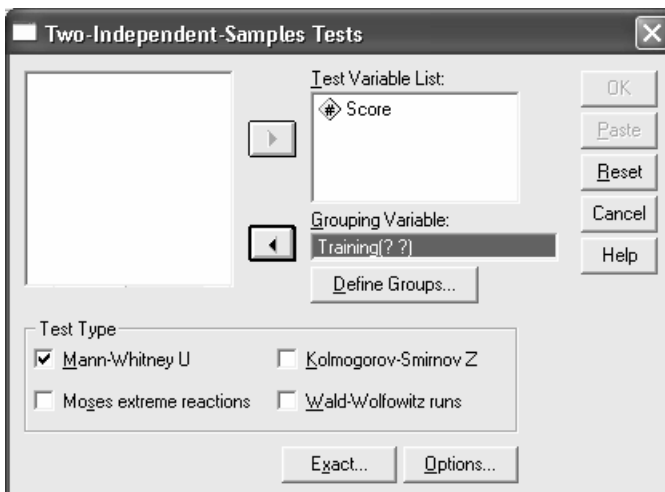
نمایش فیلم	سخنرانی
۷۰	۶۰
۹۸	۸۸
۷۹	۷۸
۸۰	۸۳
۷۵	۶۰
۷۰	۵۴
۸۷	۸۰
۷۸	۵۴
۸۹	۵۰
۸۴	۸۰
۹۶	۶۰
	۶۵

از آنجایی که پیش فرض‌های آزمون t برقرار نیست بنابراین پژوهشگر از آزمون ناپارامتریک U من- ویتنی استفاده نموده است.
برای اجرای آزمون U من- ویتنی مراحل زیر را انجام دهید:

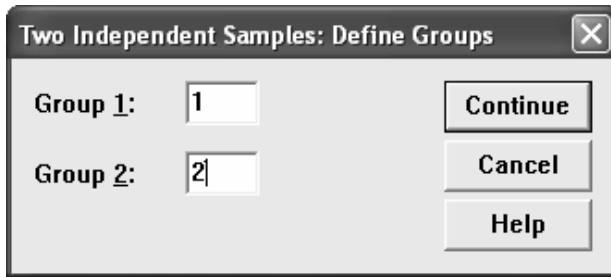
Analyze
Nonparametric test
2 Independent Samples...



بعد از اجرای این فرمان شکل زیر ظاهر می‌شود:

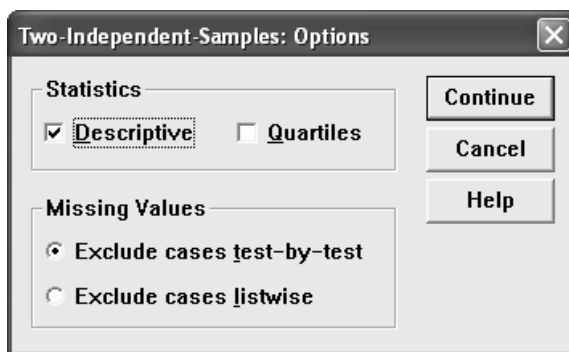


در پنجره ظاهر شده متغیر(رتبه) را به قسمت (Test Variable List) و متغیر آزمون(Examin) را به قسمت(Grouping Variable) منتقل کنید در جلوی آزمون[؟؟] قرار دارد که باید نظامیان را تعریف کرد برای تعریف دو گروه نظامیان بر روی گزینه(Define Grouping...) کلیک نمایید تا پنجره زیر آشکار شود:



در این مثال پژوهشگر چون به بررسی ۲ گروه پرداخته است در قسمت (Two Independent Samples:DefineGrouping) در قسمت (Minimum) عدد یک و در قسمت (Maximum) عدد دو قرار می دهد سپس (Continue) را کلیک کنید.

با فعال کردن گزینه (Continue) به پنجره (Two Independent Samples Tests) باز می گردید و در این قسمت برای محاسبه آمار توصیفی گزینه (Options) را کلیک نمایید تا به پنجره زیر برسید در پنجره مذکور گزینه (Descriptive) را انتخاب کنید.



سپس گزینه (Continue) را فعال نمایید تا به پنجره گفتگوی (Two Independent Samples Tests) باز گردید در پنجره گفتگوی مذکور گزینه(Mann-Whitney U) را کلیک نمایید:

Test Type

Mann-Whitney U **Kolmogorov-Smirnov Z**

Moses extreme reactions **Wald-Wolfowitz runs**

سپس (OK) را کلیک نموده تا نتایج به صورت خروجی زیر مشاهده شود:

Mann-Whitney Test

Ranks

	Training	N	Mean Rank	Sum of Ranks
Score	Film	11	15.41	169.50
	Lecture	12	8.88	106.50
	Total	23		

Test Statistics^b

	Score
Mann-Whitney U	28.500
Wilcoxon W	106.500
Z	-2.314
Asymp. Sig. (2-tailed)	.021
Exact Sig. [2*(1-tailed Sig.)]	.019 ^a

a. Not corrected for ties.

b. Grouping Variable: Training

جدول نخست، تعداد نمونه در هر گروه، میانگین رتبه‌ها و جمع رتبه‌ها را نشان می‌دهد. در جدول دوم نیز میزان U ، Z و سطح پوششش دو سویه آورده شده است. مقدار U محاسبه شده برابر ۲۸ است که این مقدار در سطح $0/05$ معنی‌دار نیست.

نکته: اگر حجم نمونه بزرگ‌تر از ۲۰ باشد توزیع نمونه‌گیری U تقریباً نرمال خواهد بود (با توجه به قضیه حد مرکزی با افزایش حجم نمونه، شکل توزیع نمونه‌گیری به نرمال نزدیک می‌شود) اگر مقدار U را به نمره Z معیار تبدیل کنیم، می‌توان از طریق ZU نیز

در مورد رد یا عدم رد فرضیه صفر قضاوت کنیم (شیولسون ترجمه کیامنش، ۱۳۸۲:۲۲۴). نحوه قضاوت بدین گونه می‌باشد که چنانچه سطح پوشش دوسویه بر دو تقسیم شده و مقدار آن از ۲/۵ درصد کمتر باشد، تفاوت دو گروه را نتیجه‌گیری نمایید و در صورتی که این میزان از ۲/۵ درصد بیشتر باشد، نبود تفاوت بین دو گروه را نتیجه‌گیری کنید. که در این مثال مقدار Z تقسیم بر ۲ برابر ۱/۱۵۷ است و چون این مقدار کمتر از ۲/۵ است، بنابراین نتیجه می‌گیریم که بین دو گروه تفاوت معنی‌داری وجود دارد.

آزمون‌های ناپارامتریک مربوط به دو گروه وابسته (2 Related samples)

آزمون T ویلکاکسون برای نمونه‌های وابسته (همبسته) مشابه آزمون t وابسته است. با این تفاوت که با استفاده از آزمون ناپارامتریک مربوط به دو گروه وابسته، می‌توان توزیع دومتغیر وابسته را بررسی کرد. آزمون ویلکاکسون یکی از آزمون ناپارامتریک مربوط به دو گروه وابسته است که در آزمون ویلکاکسون علاوه بر جهت تغییرات، اندازه نسبی این تغییرات (تفاوت‌ها) نیز مورد توجه است. یعنی با آن می‌توان درباره اینکه کدام یک از اجزاء از دیگری بزرگ‌تر است، قضاوت کرد.

$$z = \frac{T - \frac{n(n+1)}{4}}{\sqrt{\frac{n(n+1)(2n+1)}{24}}}$$

کاربرد

این آزمون نظیر آزمون U مان-ویتنی است، وقتی بکار برده می‌شود که داده‌ها دست‌کم در مقیاس ترتیبی و به صورت علامت‌گذاری شده (جفت‌های جور شده) باشند، توان آزمون ویلکاکسون در داده‌های پارامتری تقریباً با آزمون t استودنت وابسته متناظر است و توان آن برابر ۹۵/۵٪ آزمون t می‌باشد. همچنین اگر بخواهیم علاوه بر جهت تفاوت‌های موجود بین گروه‌های همتا، اندازه نسبی این تفاوت‌ها را نیز بدست آوریم، از این آزمون استفاده می‌کنیم. (سیگل ترجمه کریمی، ۱۳۸۰:۱۰۷)

محاسن

آزمون ویلکاکسون بیانگر این پاسخ است که کدام جزء از جزء دیگر بزرگ‌تر است و همچنین تفاوت‌ها را به ترتیب قدر مطلق آن‌ها رتبه بندی می‌کند علاوه بر آن جفت‌هایی که بین اجزای آن‌ها تفاوت بزرگ‌تری وجود دارد را نشان می‌دهند که این از امتیاز آزمون ویلکاکسون می‌باشد و قابل ذکر است که در این آزمون علاوه بر جهت تفاوت‌ها، اندازه نسبی این تفاوت‌ها نیز منظور می‌شود.

معایب

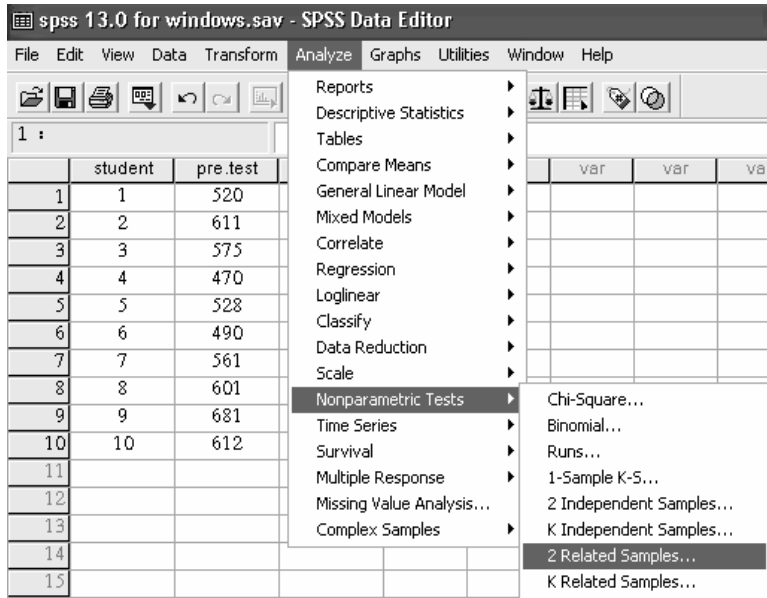
در این آزمون جفت‌هایی که نمره یکسانی دارند یعنی تفاوت آن‌ها صفر است از حجم نمونه حذف می‌شوند که این از معایب می‌باشد.

مثال: معلم یک کلاس، یک آزمون استاندارد شده را پیش از شروع تدریس در کلاس اجرا می‌کند. در پایان دوره آموزش نیز همان آزمون را درباره همان دانش‌آموزان اجرا می‌کند. هدف وی از تحقیق این است که میزان افزایش آموخته‌های شاگردان را در دوره آموزش تعیین کند که نتایج دو آزمون پیش آزمون و پس آزمون در جدول زیر آمده است:

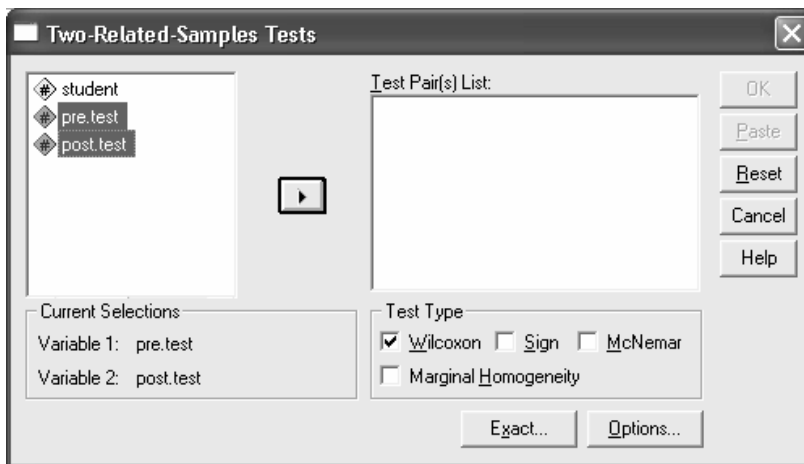
پس آزمون	پیش آزمون	دانش آموز
۵۳۰	۵۲۰	۱
۶۵۰	۶۱۱	۲
۶۰۲	۵۷۵	۳
۴۶۸	۴۷۰	۴
۵۳۱	۵۲۸	۵
۵۲۵	۴۹۰	۶
۵۷۷	۵۶۱	۷
۶۱۲	۶۰۱	۸
۶۹۸	۶۸۱	۹
۵۹۸	۶۱۲	۱۰

مراحل اجرای آزمون ویلکاکسون در Spss به شرح زیر می‌باشد:

Analyze
Nonparametric test
2 Related Samples



بعد از اجرای این فرمان شکل زیر ظاهر می شود:



در پنجره ظاهر شده متغیرها را به قسمت (Test Pair(s) List) منتقل کنید و گزینه (Wilcoxon) را انتخاب نمایید و سپس (OK) را کلیک کنید و نتایج به صورت خروجی زیر ظاهر خواهد شد:

NPar Tests
Wilcoxon Signed Ranks Test

Ranks

		N	Mean Rank	Sum of Ranks
post.test – pre.test	Negative Ranks	2 ^a	3.00	6.00
	Positive Ranks	8 ^b	6.13	49.00
	Ties	0 ^c		
	Total	10		

- a. post.test < pre.test
- b. post.test > pre.test
- c. post.test = pre.test

Test Statistics^b

	post.test – pre.test
Z	-2.191 ^a
Asymp. Sig. (2-tailed)	.028

- a. Based on negative ranks.
- b. Wilcoxon Signed Ranks Test

در جدول نخست تعداد و میانگین رتبه افرادی که نمرهای آنان در دو وضعیت افزایش، کاهش و یا ثابت مانده است آورده شده است.

- a. post.test < pre.test
- b. post.test > pre.test
- c. post.test = pre.test

(a) بیانگر این است که دو نفر از دانش‌آموزان نمره پیش‌آزمون آن‌ها بزرگ‌تر از پس‌آزمون است.

(b) بیانگر این است که هشت نفر از دانش‌آموزان نمره پیش‌آزمون آن‌ها کوچک‌تر از پس‌آزمون است.

(c) بیانگر این است که هیچ‌کدام از دانش‌آموزان نمره پیش‌آزمون و پس‌آزمون آن‌ها مساوی نیست.

در جدول دوم مقدار z و سطح پوشش آماره آزمون محاسبه شده است و چنانکه قدر مطلق z محاسبه شده بیشتر از ۱٫۹۶ باشد تفاوت معنی‌دار در پیش آزمون و پس آزمون را نتیجه‌گیری نمایید و چنانچه این مقدار کمتر از ۱٫۹۶ باشد، تساوی دو وضعیت را نتیجه‌گیری کنید. از روی مقدار sig نیز می‌توان درباره معنی‌دار بودن تفاوت‌ها اظهار نظر کرد که در این مثال مقدار sig برابر ۰٫۰۲۸ است که این مقدار در سطح ۰٫۰۵ خطا معنی‌دار است و با ۹۵٪ اطمینان می‌توان گفت که بین نمرات پیش آزمون و پس آزمون تفاوت معنی‌دار وجود دارد.

آزمون‌های ناپارامتریک مربوط به k گروه مستقل (K Independent samples)

آزمون‌های تحلیل واریانس یک متغیری داده‌های رتبه‌بندی کروسکال - والیس (H یا T) یکی از آزمون‌های ناپارامتری مربوط به K گروه مستقل است که متمم آزمون U من-ویتی می‌باشد و در مقابل آزمون F بوجود آمده و توزیع آن با χ^2 تقریباً یکی است و برای مقایسه میانگین رتبه‌های دو گروه یا بیشتر به کار می‌رود:

$$H = \frac{12}{N(N+1)} \sum_{i=1}^k \frac{R_i^2}{n_i} - 3(N+1)$$

کاربرد

این آزمون در مواردی بکار برده می‌شود که همانند آزمون F تعداد گروه‌ها دو یا بیش از دو باشد و علاوه بر آن مقیاس اندازه‌گیری دست‌کم ترتیبی باشد.

محاسن

برای مقایسه بیش از دو گروه بکار می‌رود و نیز در صورتی که فرض یکسان بودن واریانس‌ها برقرار نباشد، آنگاه می‌توان به جای آزمون F از آزمون کروسکال والیس استفاده کرد. که البته محدودیت آن نسبت به آزمون F کمتر است.

معایب

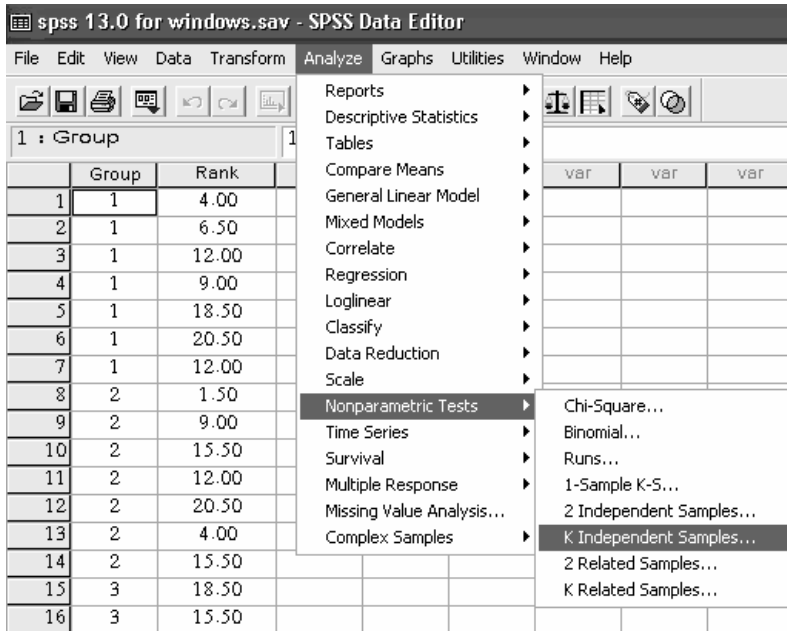
کارایی این آزمون در حدود ۹۵ درصد آزمون f است و فرضیه‌های مخالف در این آزمون همواره بدون جهت است. یعنی نشان می‌دهد که، گروه‌ها با یکدیگر تفاوت دارند. اما هرگز بیان نمی‌کند که کدام گروه بالاتر و کدام گروه پایین‌تر است. (سیگل ترجمه کریمی، ۱۳۸۰: ۲۴۲)

مثال: پژوهشگری علاقمند است که به بررسی جو خانوادگی و تأثیر آن بر پیشرفت تحصیلی درس ریاضی دانش‌آموزان ابتدایی بپردازد. به همین منظور ۲۱ نفر از دانش‌آموزان در سه گروه، خانواده‌های دارای جو دموکراسی (X)، دیکتاتوری (Y) و آزادی مطلق (Z) انتخاب و آزمون ریاضی اجرا کرد که نتایج آن به شرح زیر است:

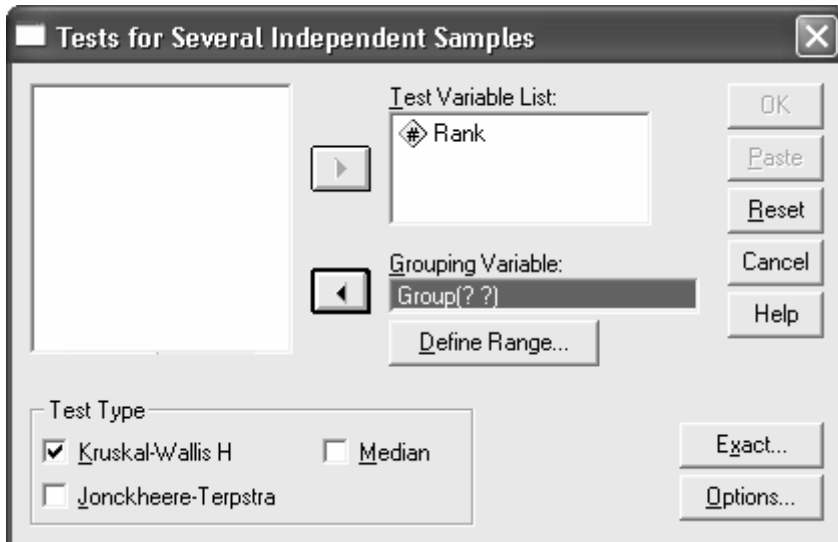
رتبه	Z	رتبه	Y	رتبه	X	دانش‌آموز
۱۸/۵	۱۸	۱/۵	۱۲	۴	۱۳	۱
۱۵/۵	۱۷	۹	۱۵	۶/۵	۱۴	۲
۴	۱۳	۱۵/۵	۱۷	۱۲	۱۶	۳
۹	۱۵	۱۲	۱۶	۹	۱۵	۴
۱۵/۵	۱۷	۲۰/۵	۱۹	۱۸/۵	۱۸	۵
۱/۵	۱۲	۴	۱۳	۲۰/۵	۱۹	۶
۶/۵	۱۴	۱۷	۱۷	۱۶	۱۶	۷

مراحل اجرای آزمون کروسکال والیس در Spss به شرح زیر می‌باشد:

Analyze
Nonparametric test
K Independent Samples...

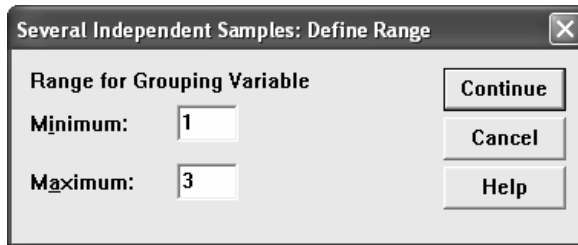


بعد از اجرای این فرمان شکل زیر ظاهر می شود:

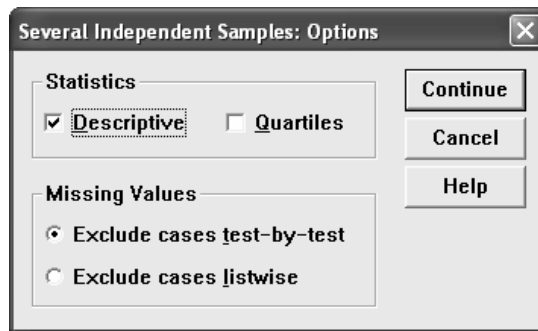


در پنجره ظاهر شده متغیر از قسمت (Test Variable List)، رتبه (Rank) را انتخاب کرده و از قسمت (Grouping Variable) (Grouping ??) را در نظر می گیریم

باید گروه را تعریف کنید. در این مثال، پژوهشگر چون به بررسی ۳ گروه پرداخته است در قسمت (Several Independent Samples: Define Range) در قسمت (Minimum) عدد یک و در قسمت (Maximum) عدد سه قرار دهید. سپس (Continue) را فعال نمائید:



تا به پنجره گفتگوی (Test for Several Independent Samples) باز گردید و چنانچه بخواهید آمار توصیفی را محاسبه کنید، گزینه (Descriptive) را کلیک نمایید، تا پنجره گفتگوی (Option...) حاصل شود و گزینه (Descriptive) را انتخاب کنید:



سپس گزینه (Continue) کلیک نمایید تا دوباره به پنجره گفتگوی (Test for Several Independent Samples) باز گردید در این پنجره گزینه (Kruskal-Wallis H) را انتخاب نمایید:



و سپس (OK) را کلیک نمایید تا نتایج به صورت خروجی زیر مشاهده گردد:

NPar Tests

Kruskal-Wallis Test

Ranks

	Group	N	Mean Rank
Rank	X	7	11.79
	Y	7	11.14
	Z	7	10.07
	Total	21	

Test Statistics^{a,b}

	Rank
Chi-Square	.277
df	2
Asymp. Sig.	.870

a. Kruskal Wallis Test

b. Grouping Variable: Group

در جدول مربوط به آزمون کروسکال والیس همانطور که مشاهده می‌کنید میانگین گروه‌ها به ترتیب از دانشکده اول تا سوم، ۱۱٫۷۹ و ۱۱٫۱۴ و ۱۰٫۰۷ می‌باشد. همچنین مقدار به دست آمده در جدول که برابر ۰٫۸۷۰ است که این مقدار از مقدار جدول کوچک‌تر است. بنابراین می‌توان نتیجه گرفت، نمرات دانش‌آموزان گروه‌های مختلف (دانش‌آموزان دارای جو خانوادگی دموکراسی، دانش‌آموزان دارای جو خانوادگی دیکتاتوری و دانش‌آموزان دارای جو خانوادگی آزادی مطلق) در درس ریاضی با یکدیگر تفاوت معنی‌داری ندارند.

مقایسه‌های بعد از تجربه زوجی (یک به یک)

همان‌طور که معنی‌دار بودن آزمون F در ANOVA یک متغیری، نشان می‌دهد که میان میانگین K جامعه تفاوت معنی‌دار وجود دارد، معنی‌دار بودن آزمون H در

KWANOVA نیز نشان می‌دهد که، تفاوت معنی‌داری بین K جامعه وجود دارد. اگر فرضیه صفر را رد کنیم، قدم بعدی آن است که تفاوت میان میانگین رتبه‌ها را به صورت آماری بررسی کنیم. در اینجا شیوه مقایسه بعد از تجربه زوجی ارائه می‌شود. روش همان روش HSD توکی است که قبلاً مورد بحث قرار گرفته است. در KWANOVA از آزمون توکی برای مقایسه میانگین رتبه‌ها استفاده می‌شود. (شیولسون ترجمه کیامش، ۱۳۸۲: ۲۴۹).

آزمون‌های ناپارامتریک مربوط به K گروه وابسته (K Related samples)

آزمون‌های ناپارامتریک مربوط به K گروه وابسته به مقایسه دو یا چند متغیر وابسته را می‌پردازد که از بین این آزمون‌ها، آزمون فریدمن مورد بررسی قرار می‌گیرد. آزمون فریدمن، یکی از آزمون‌های ناپارامتری است که در مقابل آزمون F بوجود آمده و مشابه آزمون کروسکال والیس است، با این شرط که k گروه نمونه جور شده باشند. این آزمون معلوم می‌کند که آیا میانگین رتبه‌ها یا حاصل جمع رتبه‌ها به طور معنی‌داری با یکدیگر تفاوت دارند یا خیر؟

$$T = \frac{(k-1) \left[\sum_{j=1}^k R_j^2 - n^2(k)(k+1)/4 \right]}{\frac{nk(k+1)(2k+1)}{6} - \frac{nk(k+1)^2}{4}}$$

کاربرد

این آزمون زمانی بکار برده می‌شود که مقیاس اندازه‌گیری دست کم در سطح مقیاس ترتیبی باشد. بنابراین وقتی که داده‌ها به صورت رتبه‌ای باشند آزمون فریدمن بر آزمون f ترجیح داده می‌شود. همچنین وقتی این آزمون را به کار می‌بریم که حجم گروه‌ها تقریباً یکسان باشد.

محاسن

۱. توان آزمون فریدمن برابر با قوی‌ترین آزمون پارامتری برای K گروه مستقل (یعنی آزمون F) است.

۲. یکی از شرایط اصلی کاربرد آزمون f ، همگنی واریانس است که در آزمون فریدمن نیاز به رعایت این شرط نیست.
۳. آزمون فریدمن را می‌توان با گروه‌هایی که تعداد آزمودنی‌ها خیلی کوچک است به کار برد. (سیگل، ترجمه کریمی، ۱۳۸۰:۲۱۰)

معایب

مشکل آزمون فریدمن این است که هر یک از گروه‌ها حجم‌شان باید تقریباً یکی باشد، یعنی تعداد افراد و نمونه‌ها در گروه‌های مختلف با هم برابر باشند.

مثال: پژوهشگری علاقمند است تأثیر اوقات مختلف روز (ساعت ۸ صبح، ساعت ۱۰ صبح و ساعت ۲ بعد از ظهر) را در عملکرد گروهی دانش‌آموزان در یک مدرسه مورد مطالعه قرار دهد. به همین منظور ۱۶ آزمودنی را انتخاب می‌کند و عملکرد آنان را در اوقات تعیین شده اندازه‌گیری می‌کند با استفاده از آزمون فریدمن و با احتمال ۰/۰۵ خطا تعیین کنید که آیا اوقات مختلف روز در عملکرد آزمودنی‌ها مؤثر است یا خیر؟ (حسینی، ۱۳۸۲:۱۵۸)

رتبه	۲ بعدظهر	رتبه	۱۰ صبح	رتبه	۸ صبح	دانش‌آموز
۳	۱۴	۲	۱۲	۱	۱۰	۱
۳	۱۷	۲	۱۶	۱	۱۵/۵	۲
۲	۱۲	۱	۱۰	۳	۱۴	۳
۳	۱۷	۲	۱۵	۱	۱۲	۴
۳	۱۵	۱	۱۰	۲	۱۱	۵
۱	۱۳	۲	۱۵	۳	۱۷	۶
۲	۱۸	۳	۱۹	۱	۱۸/۲۵	۷
۱	۱۴	۳	۱۷	۲	۱۷	۸
۱	۱۲	۲	۱۴	۳	۱۶	۹
۱	۹	۲	۱۰	۳	۱۱	۱۰
۳	۱۹	۲	۱۷	۱	۱۲	۱۱
۳	۱۴	۲	۱۴	۱	۱۳	۱۲
۳	۱۶	۲	۱۵/۵	۱	۱۵	۱۳
۳	۱۸	۲	۱۷	۱	۱۶	۱۴
۱	۱۱	۱	۱۴	۳	۱۶	۱۵
۲	۱۰/۵	۱	۱۰	۳	۱۶	۱۶

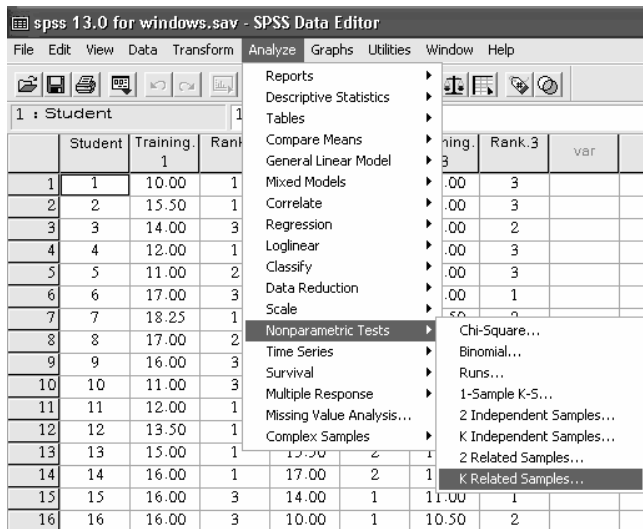
آزمون‌های آماری ناپارامتریک ۱۴۱

اکنون با توجه به نمرات می‌توانید نمره هر فرد را نسبت به خودش در هر وضعیت به دست آورید به عنوان مثال، فرد شماره یک در سه آزمون نمرات ۱۰-۱۲-۱۴ را کسب نموده، می‌توانید رتبه یک را به نمره ۱۰ رتبه دوم را به ۱۲ و رتبه سوم را به نمره ۱۴ اختصاص دهید و برای فرد دوم با توجه به نمرات رتبه‌های ۱-۲-۳ را اختصاص دهید.

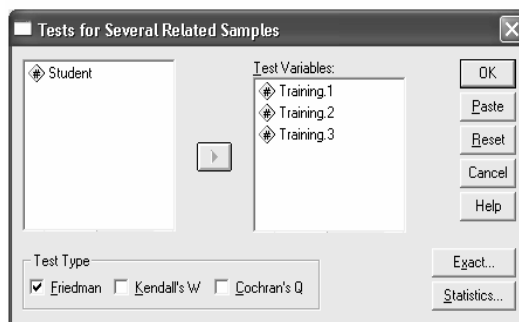
با استفاده از آزمون مناسب بیان کنید که آیا تفاوت معنی‌داری میان گروه‌ها (افراد) از لحاظ انواع روش‌های تدریس بیان شده وجود دارد یا خیر؟
برای اجرای آزمون فوق در Spss مراحل زیر را انجام دهید

Analyze

Nonparametric test
K Related Samples...



بعد از اجرای این فرمان شکل زیر ظاهر می‌شود:



در پنجره ظاهر شده متغیرها را به قسمت (Test Variable) منتقل کنید و گزینه (Friedman) را انتخاب نمایید و سپس (OK) را کلیک نموده تا نتایج به صورت خروجی زیر آشکار شود:

NPar Tests
Friedman Test

Ranks

	Mean Rank
Training.1	1.88
Training.2	1.94
Training.3	2.19

Test Statistics^a

N	16
Chi-Square	.875
df	2
Asymp. Sig.	.646

a. Friedman Test

جدول بالا مربوط به آزمون فریدمن است، در جدول مذکور به بررسی سه گروه پرداخته شده است، میانگین رتبه‌های سه گروه عبارتند از: ۱/۱،۹۱/۸۸ و ۲/۲۲ همچنین χ^2 مربوط به سه گروه با درجه آزادی ۲ برابر با ۰/۸۷۵ است که این مقدار در سطح ۰/۰۵ معنی دار نیست پس تفاوت معنی داری وجود ندارد. و عملکرد دانش‌آموزان در هر سه موقع یکسان بوده است

خودآزمایی

۱. پژوهشگری نمونه‌ای از کارکنان یک منطقه آموزش و پرورش را انتخاب و آن‌ها را در سطح آمادگی، دبستان، راهنمایی و دبیرستان تقسیم بندی نمود. سپس نظر آنان را در مورد تشکیل انجمن معلمان سؤال و به صورت جدول زیر تنظیم کرد. با یک

آزمون‌های آماری ناپارامتریک ۱۴۳

آزمون آماری مناسب و در سطح 0.01 درصد تعیین کنید بین فراوانی های جمع‌آوری شده اختلاف معنی‌داری وجود دارد یا خیر؟ (دلاور، $1380:01$)

دبیرستان	راهنمایی	دبستان	آمادگی	
۲۰	۲۵	۲۰	۶	موافق
۱۰	۱۴	۱۵	۱۶	مخالف

۲. به منظور بررسی رابطه بین سواد زنان و نگرش آنان نسبت به تنبیه فرزندان، نظر زنان باسواد و بی‌سواد را درباره ضرورت تنبیه فرزندان مورد سؤال قرار داده‌ایم. جدول زیر نظر موافق و مخالف آن‌ها را درباره تنبیه نشان می‌دهد. با یک روش آماری مناسب و در سطح 0.05 این فرضیه را که بین سواد و نگرش زنان درباره تنبیه رابطه وجود دارد، آزمون کنید.

	مخالف	موافق
باسواد	۴۵	۱۵
بی سواد	۵	۳۵

۳. از ۴۴ بیمار یک مرکز پزشکی که به‌طور تصادفی انتخاب شده بودند، پرسیده شد که از پزشک خود راضی هستند یا خیر. مؤسسه تحقیقاتی علاقه‌مند است بداند که آیا نظر بیماران در مورد پزشک معالج با شدت بیماری آن‌ها ارتباط دارد یا خیر؟ اطلاعات به‌دست آمده در جدول زیر ارائه شده است. (شیولسون ترجمه کیامنش، $1382:398$).

جمع	ناراضی	راضی	
۲۲	۱۴	۸	بیماری سخت
۲۲	۱۰	۱۲	بیماری خفیف
۴۴	۲۰	۲۰	جمع

آزمون مجذور خی را در سطح 0.05 انجام دهید.
فرضیه‌های صفر و خلاف را بنویسید.

اکنون نشان دهید که فرض‌ها برقرار هستند و شرایط لازم برای اجرای آزمون موجود است.

۴. یک گروه ۱۴ نفری از دانشجویان سال اول دانشکده به طریقی تصادفی به دو گروه گواه و آزمایشی تقسیم شدند. هدف این مطالعه، آن بود که تأثیر برنامه های مشاوره ای در مورد دانشجویان گروه آزمایشی در معدل آنها مورد بررسی قرار گیرد با استفاده از آزمون U مان-ویتی فرضیه صفر را که بین پیشرفت تحصیلی دو گروه آزمایشی و گواه تفاوت معنی داری وجود ندارد را آزمون کنید.

گروه آزمایش	گروه کنترل
۱۴	۱۷
۱۷	۱۸
۱۶	۱۴
۱۳	۱۵
۱۷	۱۴
۱۸	۱۳
۱۶	۱۱

۵. محقق می خواهد تفاوت تأثیر سه روش درمانی A,B,C را در کاهش میزان افسردگی سه گروه از بیماران افسرده که به طور تصادفی انتخاب کرده است، مورد بررسی قرار دهند. فرض کنید نمره های افراد مورد مطالعه در پایان دوره درمان به شرح زیر تعیین شده باشد. با استفاده از آزمون کروسکال - وایس فرضیه صفر را که بین روش های درمانی A,B,C در کاهش میزان افسردگی بیماران مؤثر نیست را آزمون کنید:

روش درمانی A	روش درمانی B	روش درمانی C
۱۰	۲۵	۳
۱۵	۱۷	۷
۲۰	۳۵	۱۰
۱۱	۴۰	۱۴
۱۷	۲۰	۱۵
	۱۰	۲۰
		۲۵

۶. پژوهشگری علاقمند است که به بررسی تأثیر روش تدریس حل مسأله بر پیشرفت تحصیلی دانش آموزان بپردازد به همین منظور پیش از شروع آموزش، یک پیش آزمون به عمل می آورد و سپس متغیر مستقل (روش تدریس حل مسأله ارائه کرده

آزمون‌های آماری ناپارامتریک ۱۴۵

است) از دانش‌آموزان پس از آزمون به عمل می‌آورد و با استفاده از آزمون ویلکاکسون، فرضیه صفر را در سطح 0.01 تفاوت، آزمون کنید.

پس آزمون	پیش آزمون
۱۴	۱۵
۱۷	۱۶
۱۷	۱۶/۵
۱۱	۱۰
۱۱	۱۰/۵
۱۳	۱۷
۱۲	۱۷/۲۵
۱۵	۱۱
۱۴	۱۲
۱۴	۱۳
۱۵	۱۳/۲۵
۱۴	۱۴
۱۷	۱۶

۷. فرض نمایید که نمره امتحان یک درس با سه روش تدریس (بحث گروهی، حل مساله، اکتشافی) ارزیابی شود. محقق در این زمینه بر روی یک گروه، سه روش تدریس را اجرا و سپس نتایج امتحان در هر یک از روش‌ها را یادداشت می‌کند. نتایج امتحان این ۱۶ نفر در سه آزمون در جدول زیر آورده شده است، اکنون با استفاده از آزمون فریدمن فرضیه صفر را آزمون کنید. (حسینی، ۱۳۸۲: ۱۵۷)

نمره امتحان پس از روش تدریس بحث گروهی	نمره امتحان پس از روش تدریس حل مساله	نمره امتحان پس از روش تدریس اکتشافی
۱۰	۱۲	۱۴
۱۵/۵	۱۶	۱۷
۱۴	۱۰	۱۲
۱۲	۱۵	۱۷
۱۱	۱۰	۱۵
۱۷	۱۵	۱۳
۱۸/۲۵	۱۹	۱۸/۵
۱۷	۱۷/۵	۱۴
۱۶	۱۴	۱۲
۱۱	۱۰	۹
۱۲	۱۷	۱۹
۱۳/۵	۱۴	۱۴/۵
۱۵	۱۵/۵	۱۶
۱۶	۱۷	۱۸
۱۶	۱۴	۱۱
۱۶	۱۰	۱۰/۵