آزمون مک نمار

آزمون مک نمار مانند آزمون علامت که قبلا مطرح شد، برای معنی دار بودن تغییرات به ویژه برای طرح های «قبل از/بعد از» استفاده می شود که در آن هر فرد یا آزمودنی به عنوان گواه خود به کار می رود و در آن یافته ها به صورت اسمی یا رتبه ای هستند. از این آزمون برای بررسی میزان تاثیر یک «مقاله، کتاب، سخنرانی، دوره آموزشی، ملاقات های فردی و...» استفاده می شود. مانند هنگامی که می خواهید نظر افراد را قبل از جلسه و بعد از یک جلسه سخنرانی مقایسه کنید. چنین آزمایش هایی نمونه های وابسته از داده های اسمی یا ترتیبی را در اختیار می گذارند.
در این گونه آزمون ها فرضیه های صفر و یک به صورت زیر مطرح اند. فرض صفر بیان می دارد که اختلافی بین نظرات آزمون شوندگان در قبل و بعد وجود ندارد و در مقابل فرضیه یک بین دو حالت اختلاف قائل است.

$$\begin{cases} H_0 : \mu_1 = \mu_2 \\ H_1 : \mu_1 \neq \mu_2 \end{cases}$$

برای سنجش معنی دار بودن تغییرات با استفاده از این روش باید ابتدا پاسخ های دریافت شده از آزمودنی ها را در یک جدول دو در دو به صورتی که در زیر مشاهده می کنید قرار داد. در این جدول آزمودنی هایی که تغییر را نشان می دهند در خانه های Aو Dقرار داده شده اند .Aآنهایی هستند که از + به - و Bآنهایی هستند که از – به + تغییر داشته اند. همچنین خانه های Cو Dآزمودنی هایی هستند که تغییر نکرده اند.

قبل \ بعد	+	–
+	A	B
–	C	D

از آنجایی که A+Dمجموع افرادی هستند که در آن ها تغییر صورت گرفته است. تحت فرض صفر انتظار این است که A+D/2از آزمودنی ها در یک جهت و A+D/2دیگر در جهت عکس تغییر کرده باشند. از طرفی می توان با داشتن مقادیر مشاهده شده و مقادیر مورد انتظار، شاخص کی دو را به صورت زیر تشکیل داد.

$$\sum \frac{(O_i + E_i)^2}{E_i} = \frac{(A - \frac{A+D}{2})^2}{\frac{A+D}{2}} + \frac{(D - \frac{A+D}{2})^2}{\frac{A+D}{2}} = \frac{(A-D)^2}{A+D} = \chi^2$$

عبارت $\frac{(A-D)^2}{A+D}$ دارای توزیع کی دو با یک درجه آزادی است. البته از آنجایی که یک توزیع پیوسته را برای تقریب مقادیر ناپیوسته به کار می گیریم، فرمول فوق پس از تصحیح پیوستگی به صورت زیر به دست خواهد آمد.

$$\chi^2 = \frac{(|A-D|-1)^2}{A+D}$$

آزمون مک نمار در SPSS

فرض کنید یک روانشناس کودک می خواهد چگونگی آغاز «تماس اجتماعی» در کودکان را بررسی کند. او مشاهده کرده است که کودکانی که تازه وارد مهد کودک می شوند بجای اینکه تماس اجتماعی را با کودکان برقرار کنند با بزرگسالان برقرار می کنند. او پیش بینی می کند که با افزایش میزان آشنایی و تجربه در مهد کودک، کودکان تماس اجتماعی را بیشتر با کودکان برقرار می

کنند تا با بزرگسالان. برای آزمون کردن چنین فرضیه ای او ۲۵ کودک تازه وارد به مهد کودک را مورد مشاهده قرار می دهد و آن ها را بر حسب اینکه در اولین روز «تماس اجتماعی» را با کودکان برقرار می کنند یا بزرگسالان طبقه بندی می کند. پس از گذشت یک ماه مجددا «تماس اجتماعی» آن ها را بررسی و مشاهدات را برای بار دوم ثبت می کند. نتایج مشاهدات قبل و بعد از یک ماه به صورت: ۱ =تماس با کودکان و ۲ =تماس با بزرگسالان در جدول زیر آمده است.

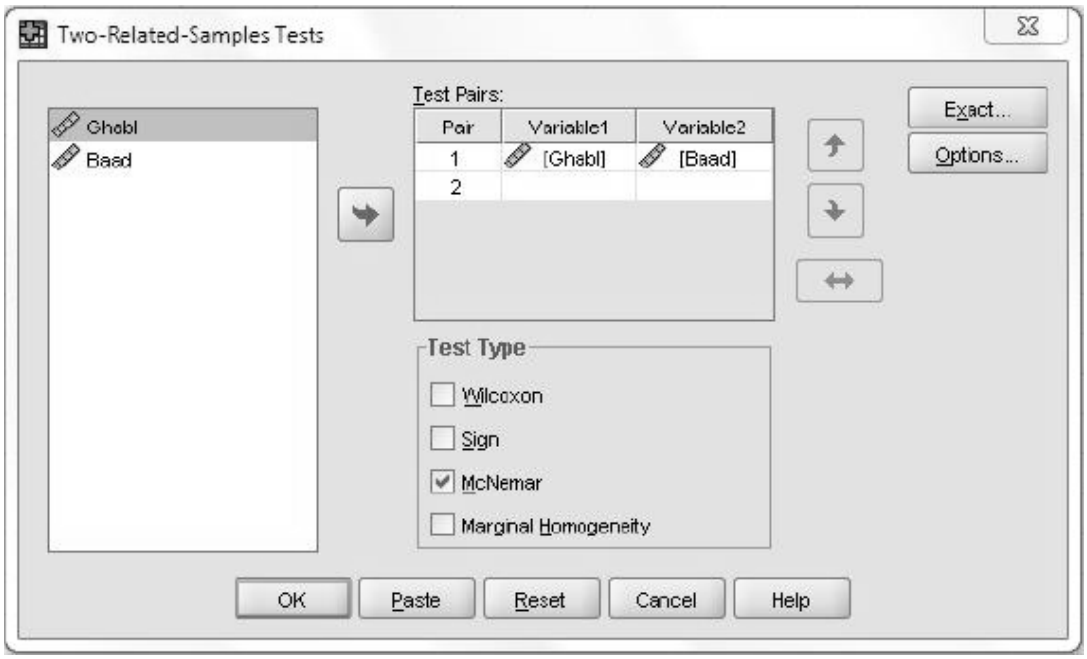
تماس قبل از دوره	۲	۲	۱	۱	۲	۲	۲	۱	۲	۱	۲	۱	۲	۲	۱	۲	۲	۲	۲	۱	۲	۲	۲	۲	۲
تماس بعد از دوره	۱	۱	۲	۱	۱	۱	۱	۲	۲	۱	۱	۲	۱	۲	۲	۱	۲	۱	۱	۱	۲	۱	۱	۱	۱

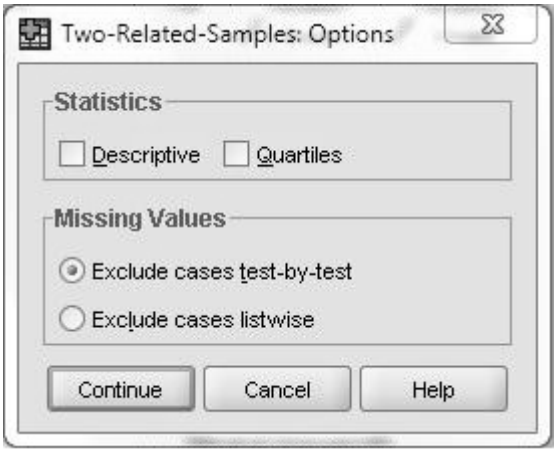
ابتدا داده ها را با تعریف دو متغیر قبل (Ghabl)و بعد (Baad)به SPSSمنتقل کنید.

فرمان زیر را اجرا کنید تا کادر محاوره آزمون های دو نمونه ای وابسته باز شود.

Analyze/Compare Means/2 Related Samples Test

در این کادر محاوره به جز آزمون مک نمار می توانید آزمون های ویلکاکسون و آزمون علامت را نیز انجام دهید. البته ما قصد داریم تنها آزمون مک نمار را برای این داده ها انجام دهیم.





در خاتمه OKرا کلیک کنید و نتیجه آزمون را ببینید.
در جدول Ghabl & Baadنتایج را به صورت خلاصه شده شاهد هستید (مانند جدولی که در ابتدای بحث به آن اشاره کردیم.)

Ghabl & Baad

Ghabl	Baad 1	Baad 2
1	3	4
2	14	4

در جدول Test Statisticsمی توانید نتیجه آزمون را مشاهده کنید. این آزمون در سطح 0/005اجرا شده است و معیار تصمیم (sig)کمتر از 0/005است فرض صفر رد می شود و مفهوم آن این است که مهد کودک در «تماس اجتماعی» کودکان تاثیر گذار است. این تاثیر بیشتر در کودکانی است که ابتدا با بزرگسالان تماس برقرار می کنند ولی بعدا به تماس با کودکان دیگر روی می آورند.

Test Statistics[b]

	Ghabl & Baad
N	25
Exact Sig. (2-tailed)	.031[a]

a. Binomial distribution used.

b. McNemar Test

آزمون من- ویتنی (Mann-Whithney)

آزمون من–ویتنی یک آزمون مقایسه ای برای مقایسه وضعیت دو گروه مستقل است و وقتی داده های یک مطالعه به صورت کیفی ترتیبی باشند بهتر است از این آزمون که یک آزمون غیر پارامتری و معادل آزمون دو نمونه مستقل t است، استفاده کرد. در این حال از آزمون t دو نمونه مستقل استفاده نمی کنیم زیرا میانگین متغیری که در مقیاس ترتیبی اندازه گیری شده باشد، به علت یکسان نبودن فاصله واحدها، معنی و مفهوم واقعی میانگین را نخواهد داشت. مثلا وقتی می خواه م ی قد دو گروه از زنان و مردان را با هم مقایسه کنیم بهتر است از مرتب کردن افراد بر حسب قد و تعیین رتبه آن ها استفاده کنیم تا مقایسه میانگین قد آن ها.

فرض کنید می خواهیم دو روش آموزش سنتی و جدید را با یکدیگر مقایسه کنیم Nدانشجو را به صورت تصادفی انتخاب و افراد نمونه را مجدداً به طور تصادفی به هر یک از دو روش اختصاص می دهیم (n_1دانشجو در روش اول و n_2دانشجو در روش دوم به طوری که ($N=n_1+n_2$و پس از پایان دوره آموزش از همه آن ها آزمون واحدی اخذ می کنیم، اینک نمرات آن ها را به ترتیب نوشته و به آن ها رتبه می دهیم. سپس مجموع رتبه های هر گروه را محاسبه کرده و به ترتیب آن ها را R_2و R_1می نامیم و در شاخص زیر قرار می دهیم.

$$W = n_1 \times n_2 + \frac{n_1(n_1+1)}{2} - R_1$$

$$W' = n_1 \times n_2 + \frac{n_2(n_2+1)}{2} - R_2$$

اگر حجم دو گروه با هم مساوی نباشند باید n_1را حجم گروه کوچک تر و n_2را حجم گروه بزرگ تر در نظر بگیرید. البته محاسبه یکی از دو Wدر بالا کافی است. چون با داشتن یکی از آن ها دیگری از رابطه زیر مشخص می شود.

$$W + W' = n_1 \times n_2$$

این آزمون یکی از قوی ترین آزمون های غیر پارامتری و جانشین مفیدی برای آزمون t دو نمونه ای مستقل، محسوب می شود. در این آزمون فرض های صفر و یک به صورت زیر هستند:

$$\begin{cases} H_0: & \text{تفاوتی بین دو گروه وجود ندارد} \\ H_1: & \text{بین دو گروه تفاوت وجود دارد} \end{cases}$$

اگر حجم نمونه ها کمتر از ۲۰ باشد باید از جدول من – ویتنی برای رد فرض صفر استفاده کرد. البته باید مقدار Wکوچک تر را در نظر گرفته و با مقدار جدول مقایسه کرد. کوچک تر بودن Wاز مقدار جدول، باعث رد فرض صفر می شود و به این معنی است که مقدار R_1از R_2کوچک تر شده و در نتیجه اختلاف بین دو گروه زیاد و فرض صفر رد می شود.

اگر حجم نمونه از ۲۰ بیشتر باشد با توجه به میانگین و واریانس Wاز شاخص زیر که دارای توزیع نرمال استاندار است استفاده کرده و مقدار آن را در سطح ۰۰۵/۰ با مقدار ۹۶/۱ جدول نرمال استاندارد مقایسه می کنیم.

$$U = Z = \frac{|W_1 - \frac{n_1 n_2}{2}|}{\sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}}$$

آزمون من- ویتنی در spss

برای مقایسه دو روش آموزش سنتی و جدید در مثال بالا دو گروه ۲۰ نفری از دانشجویان را با دو روش مختلف آموزش داده و در پایان نتیجه آزمون آن ها و رتبه آن ها را به صورت جدول زیر ثبت کرده ایم. در جدول زیر فقط کافی است نمرات دانشجویان را به spssوارد کنیم. برای این کار باید دو متغیر تعریف کنید؛ یک متغیر با عنوان Nomrehکه نمرات را در آن ثبت می کنیم و متغیر دیگر به اسم goruhکه باید گرو ه های Aو Bرا در آن تعریف کنیم.

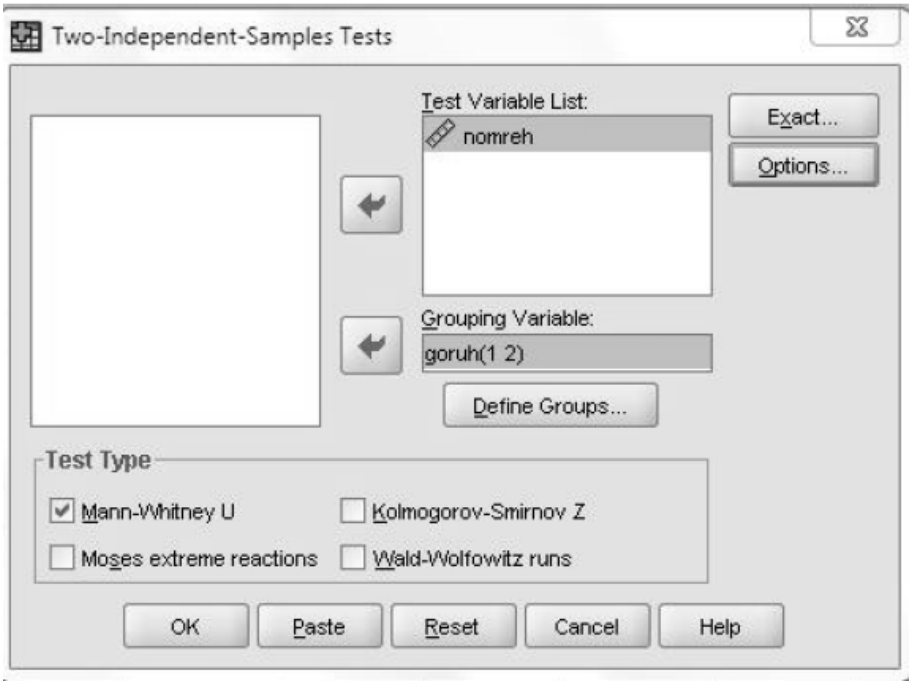
رتبه گروه B	رتبه گروه A	گروه B	گروه A
۱	۲	۵۳	۵٤
۳	٤	۵۵	۵٦
۵	۸	۵۷	٦۳
٦	۱۱/۵	۵۸	٦٦
۷	۱۱/۵	٦۲	٦٦
۹	۱٤	٦٤	٦۸
۱۰	۱۷	٦۵	۷۲
۱۳	۱۷	٦۷	۷۲
۱۵	۲۱	٦۹	۷٦
۱۷	۲۳	۷۲	۷۸
۱۹	۲۵/۵	۷٤	۸۰
۲۰	۲۷	۷۵	۸۱
۲۲	۲۹	۷۷	۸۳
۲٤	۳۰	۷۹	۸٤
۲۵/۵	۳۱	۸۰	۸۵
۲۸	۳۲	۸۲	۸۷
۳۳/۵	۳٦	۸۸	۹۰
۳۳/۵	۳۷	۸۸	۹۱
۳۵	۳۸/۵	۸۹	۹۲
۳۸/۵	٤۰	۹۲	۹٦
۳٦۵	٤۵۵		

ستون اول را در متغیر نمره وارد کرده و ستون دوم را در ادامه ۲۰ مقدار اول وارد می کنیم. داده ها پس از وارد کردن به SPSSبه صورت زیر خواهند بود:

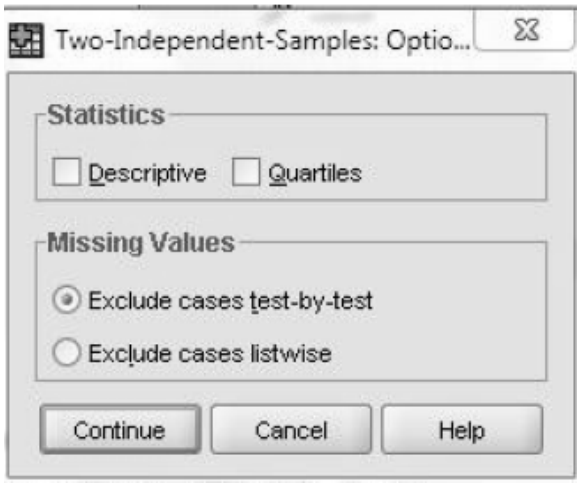
40 : goruh	2	
	nomreh	goruh	var
1	54	1	
2	56	1	
3	63	1	
4	66	1	
5	66	1	
6	68	1	
7	72	1	
8	72	1	
9	76	1	
10	78	1	
11	80	1	
12	81	1	
13	83	1	
14	84	1	
15	85	1	
16	87	1	
17	90	1	
18	91	1	
19	92	1	
20	96	1	
21	53	2	
22	55	2	
23	57	2	
24	58	2	
25	62	2	
26	64	2	

برای انجام این آزمون ابتدا باید کادر محاوره آن را از مسیر زیر باز کنید:

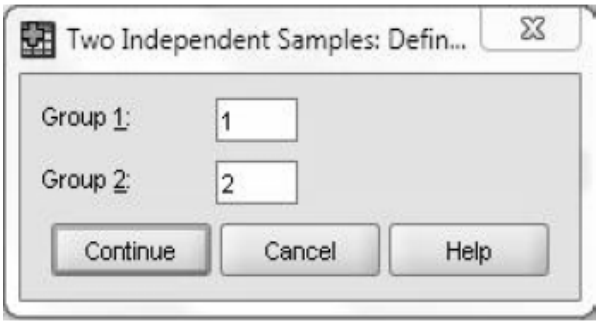
Analyze/Nonparametric Tests/2 Independent-Sampls



در این کادر محاوره متغیر نمره (nomreh)را به کادر Test Variable listو متغیر گروه بندی (goruh)را به کادر Grouping Variableمنتقل کنید. متغیر گروه بندی می تواند یک متغیر کیفی مانند جنسیت یا تأهل و مانند آن باشد.
انتخاب این متغیر بسته به آزمونی دارد که انجام می دهید. مثلاً اگر می خواهید ضریب هوشی را در دو منطقه تهران مقایسه کنید، متغیر گروه بندی، متغیر منطقه خواهد بود.
اگر متغیر گروه بندی بیش از دو گروه دارد (در این مثال منطقه) و شما می خواهید از بین آن ها دو گروه را با هم مقایسه کنید، در گزینه Define Groupsمی توانید از بین گرو ههها، دو گروه مورد نظرتان را انتخاب کنید. البته توجه داشته باشید که ممکن است متغیر گروه بندی شامل بیش از دو گروه باشد و شما بخواهید همه گرو ههها را با هم مقایسه کنید که در این صورت باید از آزمون کروسکال والیس استفاده نمایید.
از گزینه Optionمی توانید بعضی از شاخص های مهم توصیفی را برای متغیر آزمون محاسبه کنید.



گزینه Define Groupsرا کلیک کنید و هر یک از سطوح متغیر گروه بندی را در آن وارد کنید.



می توانید متغیر گروه بندی را یک متغیر کمی در نظر بگیرید. در مثال بالا اگر به جای دو منطقه بخواهید ضریب هوشی را در دو گروه سنی با یکدیگر مقایسه کنید، شما متغیر گروه بندی را یک متغیر کمی (سن) انتخاب کرده اید، می توانید برای این متغیر یک نقطه برش در قسمت Cut Pointمعرفی کنید تا آن ها را به دو گروه (کمتر از/ بیشتر از) تقسیم کرده باشید.
برای انجام این آزمون کلیدهای Continueو OKرا به ترتیب کلیک کنید و نتایج را به صورت زیر در خروجی SPSSببینید.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
nomreh	40	74.65	12.148	53	96
goruh	40	1.50	.506	1	2

در جدول بالا، میانگین و واریانس نمرات برای کل نمونه ها همراه با کم ترین و بیشترین نمره را مشاهده می کنید.
همچنین در جدول زیر میانگین رتبه ها و مجموع رتبه ها را به تفکیک برای هر گروه شاهد هستید.

Ranks

	goruh	N	Mean Rank	Sum of Ranks
nomreh	A	20	22.75	455.00
	B	20	18.25	365.00
	Total	40		

جدول اصلی، جدول Test Statisticsاست که در آن می توانید مقدار آماره ی من–ویتنی، ویلکاکسن و مقدار Zرا مشاهده کنید.
در انتهای جدول نیز مقدار معنی داری آزمون (sig)مساوی 0/231به دست آمده است که با این مقدار فرض صفر رد نمی شود یعنی دو روش آموزش تفاوت معنی داری ندارند.

Test Statistics[b]

	nomreh
Mann-Whitney U	155.000
Wilcoxon W	365.000
Z	-1.218
Asymp. Sig. (2-tailed)	.223
Exact Sig. [2*(1-tailed Sig.)]	.231[a]

a. Not corrected for ties.

b. Grouping Variable: goruh

آزمون ویلکاکسون (آزمون آماری ناپارامتری مقایسه دو گروه وابسته)

زمانی که در یک تحقیق هدف مقایسه یک متغیر در دو وضعیت متفاوت باشد و در صورت عدم برقراری فرض نرمال بودن نمونه می توان از آزمون های ناپارامتری مانند آزمون علامت، آزمون ویلکاکسون و آزمون مک نمار برای مقایسه دو گروه وابسته استفاده نمود. در این مطلب به معرفی آزمون ویلکاکسون پرداخته می شود .در آزمون ویلکاکسون نیز مانند آزمون علامت الزامی در مورد نوع توزیع متغیر مورد نظر وجود ندارد اما مقادیر متغیر مورد نظر باید پیوستگی داشته و مقیاس آن از نوع ترتیبی باشد. به عبارت دیگر اجرای این آزمون برای متغیرهایی که دارای مقوله های محدودی هستند امکان پذیر نیست.
توجه : در آزمون علامت تنها علائم مثبت و منفی به کار گرفته می شود در حالی که یک واحد ممکن است ۱ واحد اختلاف داشته باشد و واحد دیگر ۱۰ واحد. آزمون ویلکاکسون علاوه بر در نظر گرفتن مثبت یا منفی بودن داده ها، مقدار تفاوت آن ها را نیز در نظر می گیرد و چون اطلاعات بیشتری را مورد استفاده قرار می دهد نتایج دقیق تری از آزمون علامت ارائه می دهد. به مثال زیر توجه کنید:
مثال :فرض کنید نمره اطلاعات عمومی افراد در دو وضعیت قبل و بعد از شرکت در یک دوره آموزشی اندازه گیری شده باشد. می خواهیم بدانیم آیا دوره آموزشی سطح اطلاعات عمومی افراد را افزایش داده است یا خیر؟

فرد	۱	۲	۳	٤	٥	٦	۷	۸	۹
قبل از دوره آموزشی	۱٥	۱٦	۱٦.٥	۱۰	۱۰.٥	۱۷	۱۷.۲٥	۱۱	۱۲
بعد از دوره آموزشی	۱٤	۱۷	۱۷	۱۱	۱۱	۱۳	۱۲	۱٥	۱٤
تفاوت نمرات	-۱	۱	۰.٥	۱	۰.٥	-٤	-٥.۲٥	٤	۲
رتبه	٤	٤	۱.٥	٤	۱.٥	۷.٥	۹	۷.٥	٦

در این جدول ابتدا تفاوت بین نمرات حساب شده و سپس قدر مطلق این تفاوت ها مبتای محاسبه رتبه قرار داده شده است. فرد شماره ۳ و ٥ با داشتن تفاوت نمره ۰.٥ کمترین تفاوت را در بین نمونه ها دارند و بنابراین میانگین رتبه های ۱ و ۲ (۱.٥) به این دو تعلق می گیرد و به همین ترتیب سایر نمونه ها نیز رتبه بندی می شوند. حال محقق می تواند با توجه به زوج هایی که افزایش نمره و یا کاهش نمره داشته اند به جمع رتبه ها پرداخته و از آماره زیر برای آزمون مورد نظر استفاده نمود.

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

حجم نمونه مورد استفاده و T مجموع رتبه های افرادی است که افزایش نمره (یا کاهش نمره) داشته اند
آماره ی حاصل با توجه به یک طرفه یا دو طرفه بودن فرض مقابل مورد بررسی قرار می گیرد. در صورتی که حجم نمونه ها کوچک باشد این آماره با مقدار حاصل از جدول ویلکاکسون مقایسه می شود و در صورت بزرگ بودن حجم نمونه از توزیع نرمال برای رد یا پذیرش فرض صفر استفاده می نماییم.
نکته:
۱.مقدار حاصل از آماره ویلکاکسون به ازای استفاده از مجموع رتبه های مثبت یا منفی برابر می باشد و تنها تفاوت این آماره ها در علامت آن ها است. (مقدار آماره به ازای رتبه های منفی دارای علامت منفی است)
۲.در محاسبه nحجم نمونه تعداد افرادی که نمره ثابتی در دو وضعیت داشته اند از حجم کل نمونه حذف می شوند.

پیشنهادات ویژه

آزمون فریدمن

آزمون فریدمن یک آزمون ناپارامتری، معادل آنالیز واریانس با اندازه های تکراری (درون گروهی است) که از آن برای مقایسه میانگین رتبه ها در بین k متغیر (گروه) استفاده می کنیم. فرض کنید می خواهیم از یک نمونه شامل ۱۰ نفر در مورد ۵ کالا نظر خواهی کنیم. یعنی از آن ها بخواهیم که به هر یک از کالاها از نظر کیفیت امتیاز بدهند، سپس میانگین امتیازات کالاها را باهم مقایسه و بررسی کنیم که اگر اختلافات امتیازات کالاها معنی دار است کدام کالا بیشترین امتیاز و کدام کالا کمترین امتیاز را کسب کرده است. در چنین حالتی شما با ۵ متغیر روبرو هستید که این متغیر ها از لحاظ آماری به هم وابسته هستند. زیرا اندازه هایی هستند که توسط هر نمونه تکرار شده اند. تفاوت آنالیز واریانس با اندازه های تکراری (درون گروهی) با آزمون فریدمن در این است که در آنالیز واریانس شما از هر نمونه یک متغیر را به صورت تکراری در حالات مختلف اندازه گیری می گیرید. (مثلا از هر نفر (نمونه) در سه حالت ایستاده، نشسته و دراز کش فشار خون را اندازه گیری می کنید). در صورتی که در آزمون فریدمن هر یک از نمونه ها امتیازی را به چند گروه (شی یا فرد یا..) اختصاص می دهند. در هر دوی این آزمون ها متغیر ها، توسط نمونه ها مقدار گرفته اند ولی نکته مورد اختلاف این است که در آنالیز واریانس، در یک نمونه اندازه ها تکراری هستند ولی در آزمون فریدمن اندازه ها، امتیازات داده شده توسط یک نمونه است. در آزمون فریدمن فرض H مبتنی بر یکسان بودن میانگین رتبه ها در بین گروه ها است. رد شدن فرض صفر به این معنی است که در بین گروه ها حداقل دو گروه با هم اختلاف معنا داری دارند.
برای روشن شدن این موضوع و آشنایی با نحوه انجام آزمون فریدمن در نرم افزار SPSSتمرین زیر را پی می گیریم.

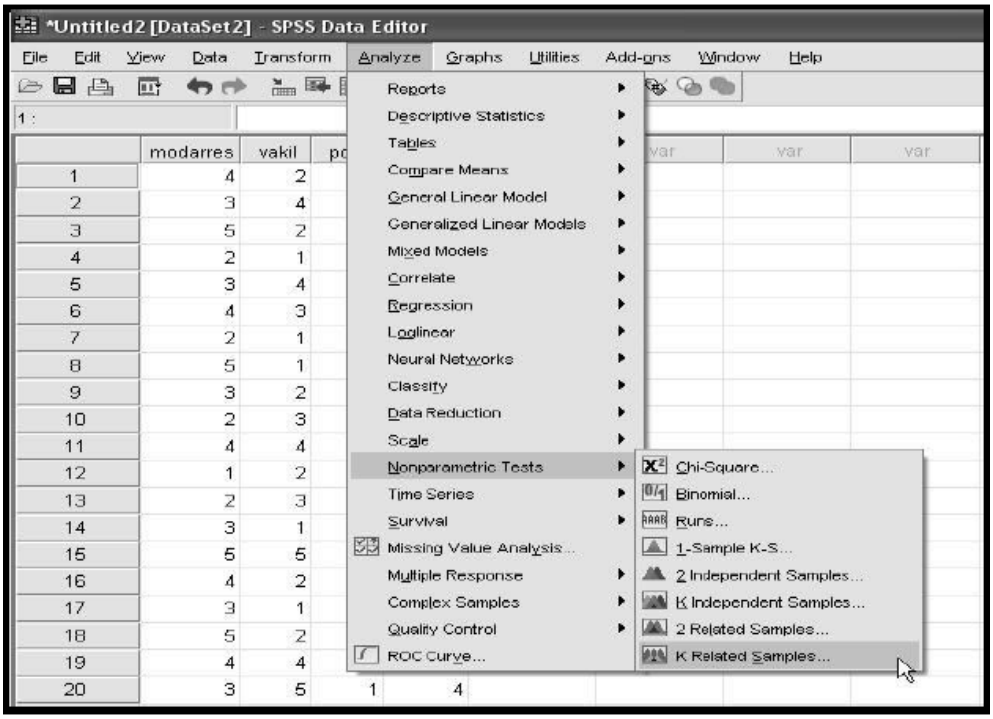
تمرین

فرض کنید از ۲۳ نفر در مورد شخصیت ۴ گروه از اقشار جامعه: مدرس، وکیل، پلیس و پزشک نظر خواهی کرده و از آن ها خواسته ایم به هر یک از شخصیت ها امتیازی بین ۱ الی ۵ اختصاص دهند. داده های این مطالعه به صورت زیر در اختیار است.
می خواهیم آزمون کنیم: آیا دیدگاه مردم نسبت به ۴ شخصیت مختلف جامعه یکسان است؟

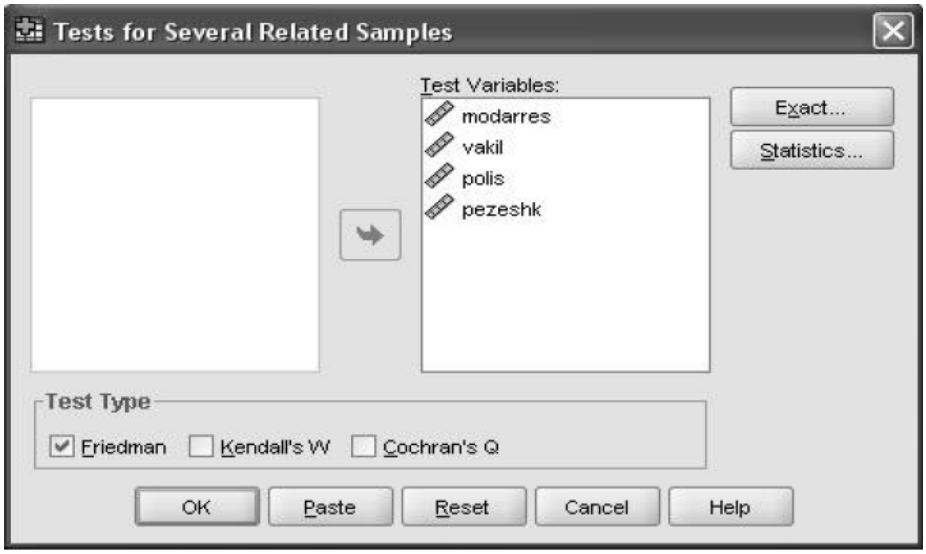
شماره نمونه	۱	۲	۳	۴	۵	۶	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	۱۴	۱۵	۱۶	۱۷	۱۸	۱۹	۲۰	۲۱	۲۲
مدرس	۳	۴	۲	۳	۵	۲	۴	۵	۳	۴	۵	۵	۵	۳	۴	۵	۳	۲	۵	۴	۵	۴
وکیل	۲	۴	۲	۱	۴	۳	۱	۱	۲	۳	۴	۲	۳	۱	۵	۲	۱	۲	۴	۵	۱	۳
پلیس	۴	۲	۱	۲	۳	۱	۲	۴	۱	۲	۳	۱	۳	۲	۲	۳	۱	۲	۱	۱	۴	۲
پزشک	۴	۳	۵	۲	۳	۴	۲	۵	۳	۲	۴	۱	۲	۳	۵	۴	۳	۵	۴	۳	۴	۳

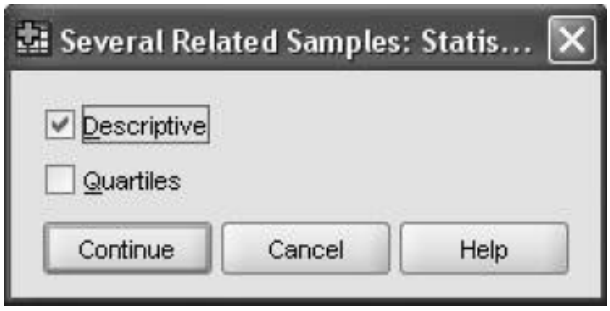
-ابتدا داده ها را در پنجره Data Editorو با ۴ متغیر، مانند جدول داده شده به spssوارد کنید.
-با توجه به ویژگی های مسئله، آزمون فریدمن برای داده های فوق آزمون مناسبی است. پس از مسیر زیر به کادر محاوره آزمون وارد شوید.

Analyze/Nonparametric Tests/K Related Samples



-در کادر محاوره آزمون متغیر های مطالعه را به چهارگوش Test Variablesمنتقل کنید.
-اگر مایل هستید برای متغیر ها آماره های توصیفی را نیز محاسبه نمایید، می توانید از گزینه Statistics مطابق شکل زیر استفاده کنید.





-توجه داشته باشید قسمت Test Typeدر کادر محاوره آزمون، گزینه مربوط به آزمون فریدمن انتخاب شده باشد.
ok-را کلیک کنید و نتیجه زیر را به صورت خروجی در spssمشاهده نمایید.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
modarres	۲۳	۳.۳۹	۱.۱۵۸	۱	۵
vakil	۲۳	۲.۵۲	۱.۳۱۰	۱	۵
polis	۲۳	۲.۱۷	۱.۰۲۹	۱	۴
pezeshk	۲۳	۳.۹۱	۱.۰۸۳	۲	۵

Ranks	
	Mean Rank
modarres	۲.۷۸
vakil	۲.۱۳
polis	۱.۶۷
pezeshk	۳.۴۱

Test Statistics[a]	
N	۲۳
Chi-Square	۲۶.۶۹۴
df	۳
Asymp. Sig.	.۰۰۰
a. Friedman Test	

نتایج و تفسیر

در جدول Ranksمیانگین رتبه های هر یک از شخصیت ها را می بینید. و جدول Test Statisticsمحتوی نتیجه اصلی آزمون است. همانگونه که در جدول شاهد هستید می توان مقدار آماره مربع کی را با ۳ درجه آزادی و همچنین سطح معنی داری آزمون P-Valueرا با مقدار صفر مشاهده کرد که نشان از رد شدن فرض H_0دارد. با توجه به خروجی های بالا نتیجه نهایی اینکه شخصیت های اقشار از نظر مردم متفاوت است. بر این اساس پزشک بالاترین امتیاز و پلیس دارای کمترین امتیاز است.

آزمون کلموگروف-اسمیرنف (K-S)

این آزمون به عنوان یک آزمون تطابق توزیع برای داده های کمی است. فرض کنید محققی نمونه ای از اندازه های کمی در اختیار دارد و می خواهد تعیین کند که آیا این نمونه از جامعه ای با توزیع نرمال بدست آمده است یا خیر؟ آزمون نرمال بودن یک توزیع یکی از شایع ترین آزمون ها برای نمونه های کوچک است که محقق به نرمال بودن آن شک دارد. برای این هدف آزمون K-S، آزمون مناسبی است. در نرم افزار spssاز این آزمون برای تطابق چهار توزیع مختلف نرمال، پواسن، نمایی و یکنواخت استفاده شده است. اساس این روش بر اختلاف بین فراوانی تجمعی نسبی مشاهدات با مقدار مورد انتظار تحت فرض صفر است. فرض صفر می گوید که نمونه انتخاب شده دارای توزیع نرمال، (پواسن، نمایی یا یکنواخت) است. آزمون کلموگروف–اسمیرنف برای تطابق توزیع، احتمال های تجمعی مقادیر در مجموعه داده هایتان را با احتمال های تجمعی همان مقادیر در یک توزیع نظری خاص مقایسه می کند. اگر اختلاف آن به قدر کافی بزرگ باشد، این آزمون نشان خواهد دادکه داده های شما با یکی از توزیع های نظری مورد نظر تطابق ندارد. در این آزمون اگر معیار تصمیم (P-Value)کمتر از 5%باشد فرض صفر رد می شود یعنی داده ها نمی توانند از یک توزیع خاص مانند نرمال، پواسن، نمایی یا یکنواخت باشند. برای درک بهتر این آزمون به یک مثال عملی که با نرم افزار spssانجام شده است توجه کنید.

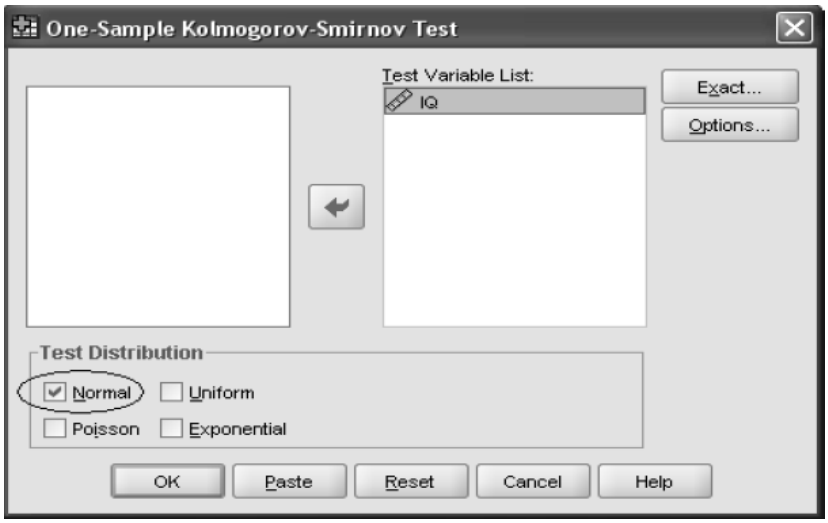
تمرین عملی (تطابق توزیع(

محققی نمره های بهره هوشی ۵۰ نفر را در اختیار دارد و برای انجام یک آزمون، فرض نرمال بودن داده ها لازم است. او می خواهد مطمئن شود که داده ها دارای توزیع نرمال هستند یا خیر؟ در جدول زیر نمره بهره هوشی این نمونه ثبت شده است:

71.7	104.6	101.1	122.6	116.5	87.7	105.9	107.5	92.4	107.3
76.3	90.5	98.6	99.4	118.6	85.7	118.5	107.2	81.8	104.4
91.5	90.7	128.7	118.8	103.8	123.1	102.7	95.3	105.1	70.7
100.3	100	117.2	135.1	111	90.8	81.9	103.2	112.1	116.9
84.4	96.4	120.7	92.1	118.4	93.7	112.4	101	88.8	104.6

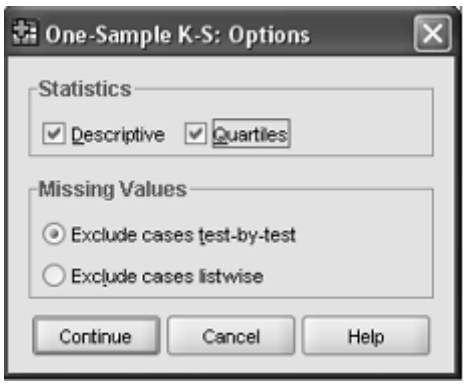
برای بررسی اینکه آیا داده ها از یک توزیع نرمال به دست آمده اند یا خیر، از آزمون کلموگروف–اسمیرنف استفاده می کنیم. در نرم افزار spssبه قسمت Variable viewبروید و ابتدا یک متغیر Scaleبه نام IQتعریف کنید و داده ها را در Data Viewوارد کنید و مراحل زیر را برای انجام این آزمون دنبال کنید:

-از فرمان Analyze/Nonparametric Test/1 SamplesK-Sبه کادر محاوره آزمون کلموگروف–اسمیرنف وارد شوید.



-متغیر IQرا به فهرست متغیرهای آزمون (Test Variable List)وارد کنید. توجه داشته باشید گزینه Normalدر قسمت Test Distributionکه به طور پیش فرض انتخاب می شود، تغییر نکرده باشد.

-گزینه Optionsرا برای محاسبه بعضی از شاخص های توصیفی انتخاب کنید.



-کلیدهای Continueو OKرا به ترتیب کلیک کنید تا آزمون انجام شود. نتایج را در جدول زیر مشاهده کنید:

One-Sample Kolmogorov-Smirnov Test

			IQ
N تعداد			50
Normal Parameters[a]	Mean	میانگین	102.1940
	Std. Deviation	انحراف معیار	14.56615
Most Extreme Differences حد اکثر اختلافات	Absolute		.077
	Positive		.058
	Negative		-.077
Kolmogorov-Smirnov Z		آماره کلموگروف-اسمیرونوف	.544
Asymp. Sig. (2-tailed)		معیار تصمیم	.928

a. Test distribution is Normal.

به جدول نتایج و معیار تصمیم (P-Value)دقت کنید می بینید که مقدار 0/928نشان از پذیرش فرض صفر دارد. یعنی دلیلی برای رد این فرضیه که "نمونه مورد نظر از توزیع نرمال به دست آمده است"، وجود ندارد. به عبارتی توزیع این نمونه، نرمال است.

آزمون کروسکال والیس

از آزمون کروسکال-والیس که یک آزمون غیر پارامتری و از سری آزمون های آنالیز واریانس محسوب می شود، برای مقایسه های سه و بیشتر از سه گروه استفاده می کنیم. روش کروسکال-والیس این فرضیه را که kگروه نمونه از یک جامعه آماری مشترک یا جامعه آماری شبیه به هم که با توجه به میانگین ها استخراج شده اند، آزمون می کند. آنالیز واریانس یک طرفه کروسکال-والیس با استفاده از رتبه ها آزمون فوق العاده مفیدی برای تصمیم گیری درباره این است که آیا kگروه نمونه مستقل از جامعه های آماری مختلف آمده اند یا نه؟ بدیهی است که نمونه ها بدون استثنا اختلافاتی با یکدیگر دارند ولی سوال این است که آیا اختلافات مشاهده شده در نمونه ها نماینده اختلافات موجود در جوامع هستند یا ناشی از شانس و تصادف اند؟ فرضیه صفر در این آزمون بر خلاف فرض مقابل آن، تاکید بر عدم اختلاف بین گروه ها دارد. این فرضیه با توجه به میانگین ها، مبنا را بر شباهت kنمونه از یک جامعه مشترک می گیرد. یعنی دو فرضیه صفر و یک به صورت زیر مطرح می شوند.

$$\begin{cases} H_0 : \mu_1 = \mu_2 = \ldots = \mu_k \\ H_1 : \mu_i \neq \mu_j \quad i \neq j \end{cases}$$

در این آزمون همه kنمونه را روی هم ریخته (تا Nمشاهده به دست آید) سپس هر یک از Nمشاهده را به صورت رتبه در می آوریم. کوچک ترین مقدار رتبه یک و بیش ترین مقدار آخرین رتبه (رتبه Nام) را به خود اختصاص می دهد. اینک برای هر یک از Kگروه، مجموع رتبه ها را محاسبه می کنیم. آزمون کروسکال والیس معلوم می کند که آیا این مجموعه رتبه ها چنان با یکدیگر تفاوت دارند که نتوان گفت آن ها از یک جامعه آماری مشترک استخراج شده اند.
شاخص آماری کروسکال والیس به صورت زیر است:

$$H = \frac{۱۲}{N(N+۱)} \sum_{j=۱}^{k} \frac{R_j^{۲}}{n_j} - ۳(N+۱)$$

که اجزای این فرمول به صورت زیر اند:
=kتعداد گروه ها
=n_jتعداد نفرات در هر گروه
=Nتعداد کل نفرات در بین همه گروه ها
=R_jمجموع رتبه ها در هر گروه

شاخص بالا برای نمونه های بزرگ (اگر تعداد نفرات در هر گروه بیشتر از ۵ نفر باشد) تقریبا دارای توزیع کی دو با k-1درجه آزادی است که با توجه به سطح معنی داری تعیین شده برای رد یا پذیرش فرض صفر می توان از جدول توزیع کی دو استفاده کرد. ولی اگر تعداد نفرات در هر نمونه کمتر از ۵ و تعداد گروه ها هم ۳ تا باشند باید از جدول مربوط به آزمون کروسکال-والیس استفاده کرد.

آزمون کروسکال-والیس در spss

فرض کنید یک محقق تعلیم و تربیت می خواهد میزان اقتدار طلبی را در بین اقشار مختلف با یکدیگر مقایسه کند. او ۵۴ نفر از ۵ قشر مختلف از مردم را انتخاب و تست اقتدار طلبی را از آنان به عمل آورده است. نتایج را در جدولی که در ادامه آمده است مشاهده نمایید.

این داده ها را با تعریف دو متغیر به spssوارد کنید. یک متغیر به نام Azmunکه نمرات اقتدارطلبی را برای تمام گروه ها و در امتداد هم در آن وارد می کنید و دیگری به نام Goruhکه شماره گروه ها را باید در آن ثبت نمایید.

گروه ۱	گروه ۲	گروه ۳	گروه ٤	گروه ٥
۹٥	۱۰۸	۱۳۰	۱۲٥	۱۱۲
۹۸	۱۲٤	۱۳٤	۱۲۷	۱۱٥
۱۰۲	۱۰۷	۱۲۸	۱۱۸	۱٤۰
۱۲۰	۱۰٦	۱۲۲	۱۰۷	۱٤۱
۱۱٦	۱۱۳	۱۳۷	۱۰۰	۱۲۰
۱۰۸	۱۲٥	۱٤۰	۱۰۹	۱۰۱
۱۰۹	۱۲٦	۱٤٥	۱۲۰	۹٥
۱۰۰	۱۳٦	۱٤۱	۱۱۱	۱۰۹
۹۹	۱۳٥	۱٤۸	۱۳۱	۱۰۱
۹٤	۱۲٤		۱۳٤	۱۱٦
۱۰۲	۱۲٤			۱۲٥
	۱۱۰			۱۳۱

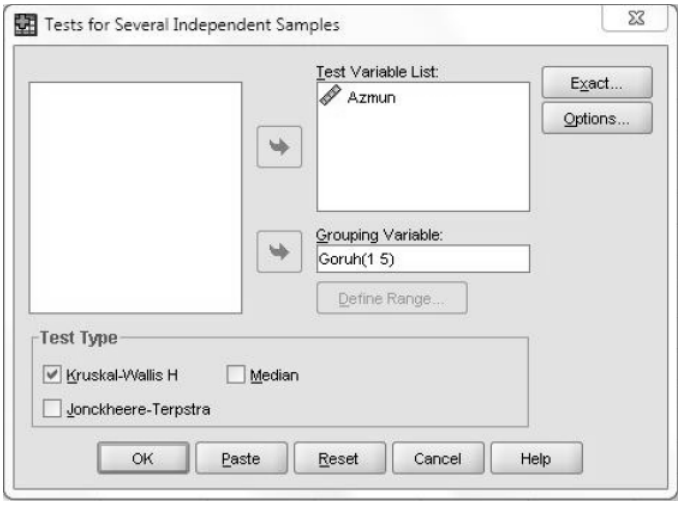
به تصویر زیر که داده ها را در spssنشان می دهد، دقت کنید.

54 : Azmun | 131

	Azmun	Goruh	var
1	95	1	
2	98	1	
3	102	1	
4	120	1	
5	116	1	
6	108	1	
7	109	1	
8	100	1	
9	99	1	
10	94	1	
11	102	1	
12	108	2	
13	124	2	
14	107	2	
15	106	2	
16	113	2	
17	125	2	
18	126	2	
19	136	2	
20	135	2	
21	124	2	
22	124	2	
23	110	2	
24	130	3	
25	134	3	
26	128	3	
27	122	3	

-با توجه به ویژگی های مسئله، آزمون کروسکال-والیس برای داده های فوق آزمون مناسبی است. پس از مسیر زیر به کادر محاوره آزمون وارد شوید.

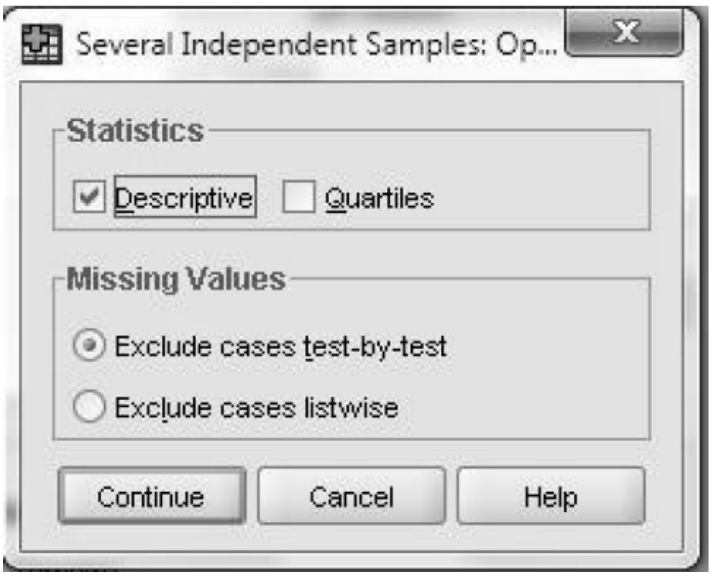
Analyze/Nonparametric Tests/K Independent Samples



-در کادر محاوره آزمون متغیرهای مطالعه را به چهارگوش Test Variables list منتقل کنید.

-توجه داشته باشید که در قسمت Test Type در کادر محاوره آزمون، گزینه مربوط به آزمون کروسکال-والیس انتخاب شده باشد.

-اگر مایل هستید برای نمرات آزمون، بعضی شاخص های توصیفی را محاسبه نمایید، از گزینه Options استفاده کنید و در کادر محاوره باز شده مانند شکل زیر گزینه Descriptive را انتخاب کنید.



-به ترتیب Continue و OK را کلیک کنید و نتیجه آزمون را به صورت زیر در خروجی spss مشاهده نمایید. در جدول Descriptive Statistics آماره های میانگن، انحراف معیار، مینیمم و ماکزیمم نمرات را مشاهده می کنید.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum
Azmun	54	118.41	14.598	94	148
Goruh	54	3.00	1.467	1	5

Ranks

	Goruh	N	Mean Rank
Azmun	1	11	11.64
	2	12	29.17
	3	9	45.83
	4	10	27.70
	5	12	26.46
	Total	54	

در جدول Ranks تعداد افراد در هر گروه و میانگین رتبه های هر یک از اقشار را می بینید.

Test Statistics[a,b]

	Azmun
Chi-Square	23.620
df	4
Asymp. Sig.	.000

a. Kruskal Wallis Test
b. Grouping Variable: Goruh

و جدول Test Statistics محتوی نتیجه اصلی آزمون است. همان گونه که در این جدول شاهد هستید می توان مقدار آماره مربع کی را با ۴ درجه آزادی و همچنین سطح معنی داری آزمون P-Value را با مقدار صفر مشاهده کرد که نشان از رد را با مقدار صفر مشاهده کرد که نشان از رد H_0 دارد. با توجه به خروجی بالا، نتیجه نهایی این است که اقتدار طلبی در بین اقشار مختلف متفاوت است.

آزمون تصادفی بودن

اگر بنا باشد که محققی بتواند از نتایج به دست آمده از یک نمونه، به نتایجی درباره جامعه برسد، باید نمونه او یک نمونه کاملا تصادفی باشد. یک آزمایش دو حالتی مانند پرتاب یک سکه را در نظر بگیرید. اگر چند بار پی در پی سکه را پرتاب کنید و نتیجه آزمایش را ثبت نمایید، به طور یقین با یک دنباله از نتایج تصادفی روبرو خواهید بود. یا مثلا اگر جنسیت مراجعه کنندگان به یک فروشگاه را به صورت دنباله‌ای از نتایج ثبت کنید، داده‌های حاصل، تصادفی خواهند بود. آزمون تصادفی بودن را هنگامی به کار بگیرید که می خواهید از تصادفی بودن توالی مقادیر متغیرها مطمئن شوید. به عبارت دیگر برای اینکه بخواهید در یک نمونه که در آن رویدادهای مختلف ثبت شده است، تصادفی بودن ثبت رویدادها را آزمایش کنید، از این آزمون استفاده کنید. در این آزمون مقادیر به دست آمده با یک مقدار معلوم (که در پیش فرض spssاین مقدار، میانه در نظر گرفته می‌شود) مقایسه شده و مقادیر کوچکتر و بزرگتر از آن تشکیل دهنده تعداد دورها خواهند بود. منظور از دور، تغییر از یک حالت به حالت دیگر است. مثلا در دنباله زیر اگر علامت مثبت را یک حالت و علامت منفی را حالت دیگر در نظر بگیرید، تعداد دورها ۷ است زیرا:

+ + + + − − + + − − − + − − − − + +

۷ ٦ ٥ ٤ ٣ ٢ ١

اگر تعداد دورها خیلی بزرگ یا خیلی کوچک باشند، نشانه عدم تصادفی بودن داده ها است. در این آزمون فرض بر تصادفی بودن داده ها استوار است و اگر فرض رد شود، تصادفی بودن داده ها به مخاطره می افتد و مفهوم آن این است که داده ها گزینشی بوده اند. فرض های صفر و یک به صورت زیر مطرح اند.

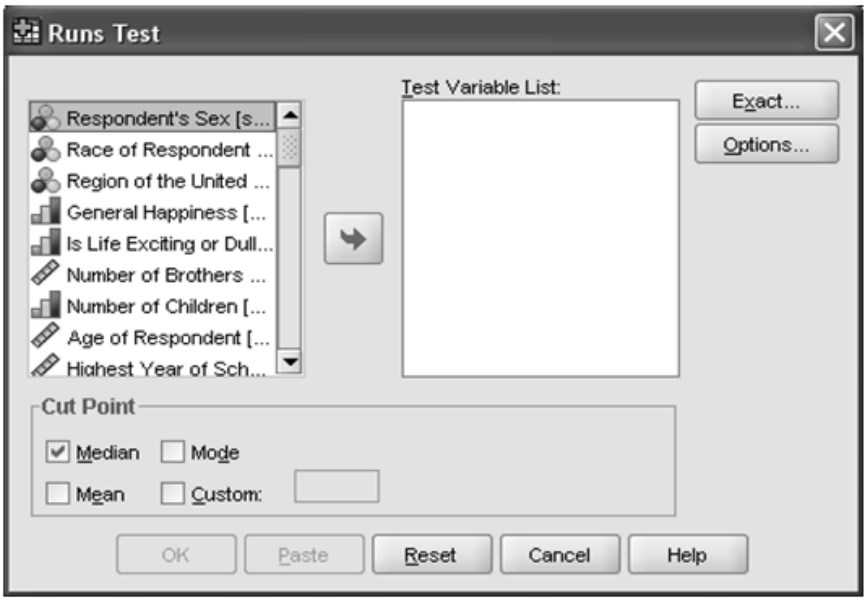
H_0 : داده ها کاملا تصادفی هستند

H_1 : داده ها تصادفی نیستند

آزمون تصادفی بودن (Runs Test)در spss

برای آزمون تصادفی بودن مقادیر یک متغیر از فرمان Runs Testاستفاده کنید تا کادر محاوره آن مانند کادر محاوره زیر باز شود.

Analyze/Nonparametric Test/Runs



متغیر مورد نظر را به چهار گوش Test Variable List منتقل کنید.

در بخش Cut point یکی از گزینه های(Median) ؛ نما (Mod)، میانگین (Mean)را انتخاب کنید. در قسمت Customمی توانید مقدار دل خواهی به عنوان Cut Pointدر نظر بگیرید. از گزینه Optionsمی توانید بعضی از شاخص های توصیفی را محاسبه کنید.

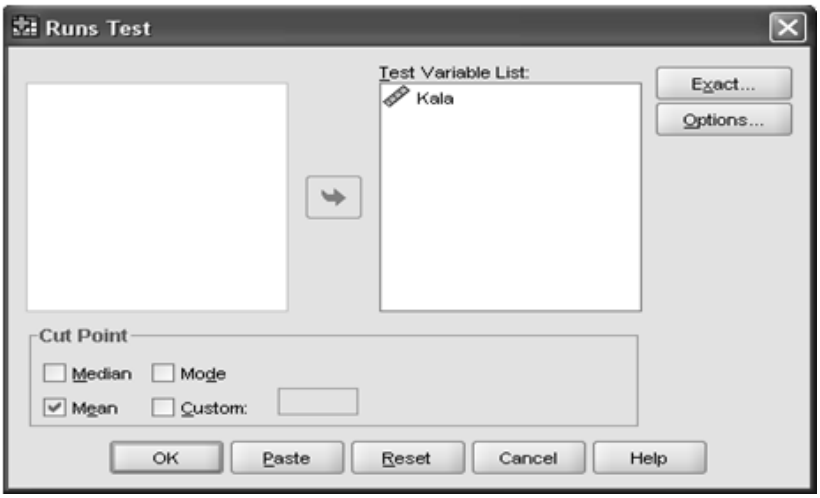
تمرین عملی

نظر ۲۰ نفر در مورد سه کالا که با کد های ۱ و ۲ و ۳ مشخص شده اند، خواسته شده و نتایج آن در جدول زیر ثبت شده است.

شماره نمونه	۱	۲	۳	٤	٥	٦	۷	۸	۹	۱۰	۱۱	۱۲	۱۳	١٤	١٥	١٦	١٧	١٨	١٩	٢٠
انتخاب کالا	۳	۱	۲	۱	۲	۳	۱	۲	۳	۱	۲	۱	۱	۳	۲	۳	۲	۳	۲	۱

آزمون کنید آیا دنباله نتایج به دست آمده به صورت تصادفی جمع آوری شده است؟
داده ها را به ترتیبی که در جدول بالا آمده است، با یک متغیر به نام (Kala)در Data Viewوارد کنید. دقت کنید ترتیب وارد کردن داده ها را تغییر ندهید، زیرا در نتیجه آزمون تاثیر منفی خواهد گذاشت. در این آزمون می خواهیم بررسی کنیم نحوه انتخاب افراد در نظر خواهی به صورت تصادفی و نه گزینشی بوده است. برای این منظور آزمون تصادفی بودن (Runs)را از فرمان زیر فراخوانی می کنیم:

Analyze/Nonparametric Test/Runs



در کادر محاوره باز شده، متغیر kalaرا به چهار گوش Test Variable Listمنتقل کنید. و Quartilsرا علامت دار کنید و okرا کلیک نمایید. نتایج در خروجي به صورت جدول هاي زیر خواهند بود.

Descriptive Statistics

	N	Mean	Std. Deviation	Minimum	Maximum	Percentiles		
						25th	50th (Median)	75th
Kala	20	1.95	.826	1	3	1.00	2.00	3.00

Runs Test

	Kala
Test Value[a]	1.95
Cases < Test Value	7
Cases >= Test Value	13
Total Cases	20
Number of Runs	12
Z	.711
Asymp. Sig. (2-tailed)	.477

a. Mean

از آنجایی که Test Valueرا میانگین انتخاب کرده ایم و مقدار آن 95/1به دست آمده است، تعداد نمونه های کمتر از میانگین 7و تعداد نمونه های بیشتر یا مساوی با میانگین 13به دست آمده است که در جدول Runs Testمشاهده می کنید. تعداد دورهای گردش Number of Case، 12به دست آمده است و نتیجه پایانی اینکه با توجه به سطح معنی داری (Asymp. Sig. (2-tailed)=0477که بیشتر از 5درصد است، فرض صفر رد نمی شود. یعنی تصادفی بودن این داده ها پذیرفته است.

آزمون دو جمله ای (آزمون نسبت یک نمونه ای)

آزمون دو جملهای یک آزمون تطابق توزیع برای دادههای اسمی است. البته در نرم افزار spssمی توان این آزمون را برای هر متغیری که تنها شامل دو مقدار باشد، انجام داد. یک آزمون دو جمله ای، توزیع یک یا چند متغیر دوحالتی را با توزیع دو جملهای و با یک احتمال مشخص مقایسه می‌کند. اگر متغیر بیش از دو حالت دارد، می توانید از آزمون «کی دو» استفاده کنید.
به عنوان مثال پزشکی ادعا می‌کند که برای یک بیماری خاص، روش درمان بهتری دارد. زیرا از ده بیماری که معالجه کرده است ۷ نفر آن ها بهبود پیدا کرده‌اند. در صورتی که درمان هائی که تاکنون صورت گرفته بیشتر از ۵۰٪ موفقیت نداشته است. سئوال این است که آیا واقعاً روش درمانی ادعا شده پزشک بهتر از روش‌های قبل است؟ برای پاسخ دادن به این پرسش که آیا نسبت موفقیت در یک آزمایش دو حالتی با یک مقدار خاص برابری دارد یا نه؟ از آزمون دو جملهای استفاده می‌کنیم. آزمون‌های دو جملهای که به آزمون نسبت نیز موسوم هستند، به آزمایش‌هایی تعلق دارد که دارای خصوصیات زیر باشند:
۱-هر یک از آزمایشات فقط دو حالت داشته باشند مانند «بهبودی/عدم بهبودی» «استاندارد/غیر استاندارد» «موفقیت/شکست» «سالم/معیوب» و مانند این ها. چنین آزمایش‌هایی را که تنها منجر به دو پیامد می‌شوند، آزمایش‌های برنولی می‌گویند.
توجه داشته باشید که می توان با تعیین یک نقطه در متغیرهای کمّی، آنها را به دو گروه کمتر/ بیشتر (از آن نقطه) تفکیک کرد.
۲-تمام مشاهدات باید مستقل باشند. منظور این است که پاسخ یک نمونه بر روی پاسخ نمونه دیگری تأثیر گذار نباشد.
در حالت کلی فرضیه آزمون دو جملهای به این صورت مطرح است که:

$$\begin{cases} H_0 : p = p_0 \\ H_1 : p \neq p_0 \end{cases}$$

که در مثال بالا فرضیه ها به صورت زیر است:

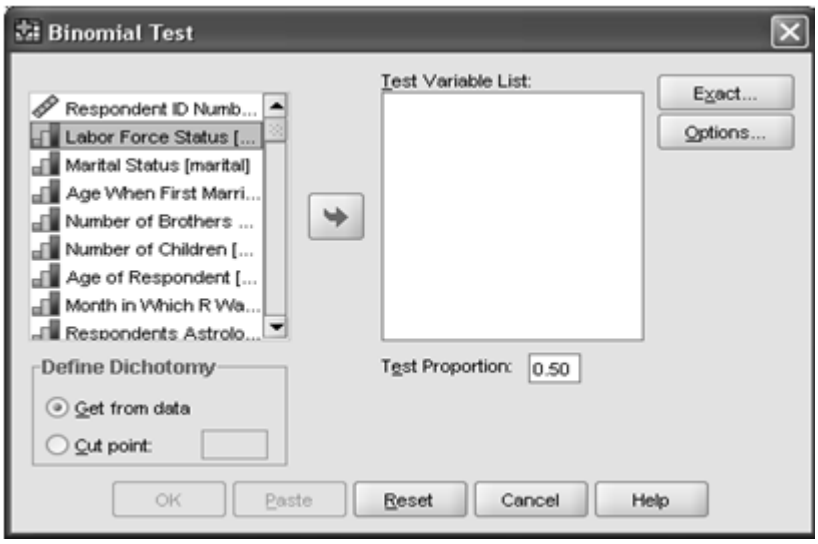
$$\begin{cases} H_0 : p = ۰/۵ \\ H_1 : p \neq ۰/۵ \end{cases}$$

تحت فرض H_0نسبت هر یک از حالت ها با هم یکسان و برابر 5/0است. در مثال بالا چون نسبت مشاهده شده در نمونه 7/0بدست آمده است، با این آزمون می‌خواهیم مطمئن باشیم که نتیجه بدست آمده ناشی از شانس و تصادف نیست بلکه واقعاً 70درصد در مقابل 50درصد نشان از روش مناسبی است که پزشک به کار برده است.
برای تعیین آنکه آیا میزان بهبودی مشاهده شده 70درصد (در صورتی که میزان واقعی ۵۰درصد است)، محتمل است یا خیر؟ از آزمون دو جمله‌ای استفاده می‌کنیم.

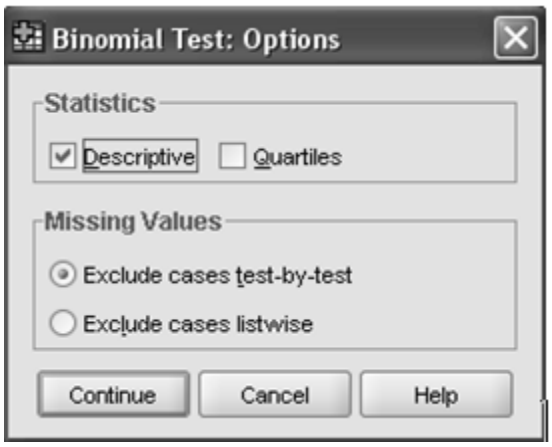
آزمون دوجمله ای در spss

برای دسترسی به این آزمون در نرم افزار spssباید از دستور زیر استفاده کنید:

Analyze/ Nonparametric Test/ Binomial Test



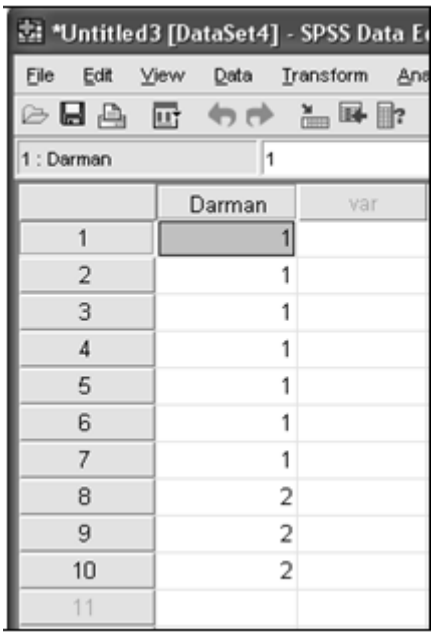
-در پنجره بازشده یک یا چند متغیر انتخاب کرده و به Test Variable Listمنتقل کنید.
-اگر این متغیرها دو حالتی هستند از گزینه Get From Dataاستفاده کنید و اگر دو حالتی نیستند و می‌خواهید آنها را دو حالتی کنید، گزینه Cut Pointرا بکار گیرید و مقداری را در آن وارد کنید. مقادیری از داده‌ها که زیر نقطه Cut Pointهستند، یک گروه تشکیل می‌دهند و مقادیری که بیشتر یا -مساوی نقطه Cut Pointمی‌باشند، گروه دوم را تشکیل می‌دهند.
-در قسمت Test Protectionنسبت آزمون را مشخص کنید.
-گزینه Optionsرا انتخاب کنید.



را برای محاسبه بعضی از شاخص‌های توصیفی ضروری و گزینه Quartilseرا برای محاسبه چارک ها انتخاب کنید.
-OKرا کلیک کنید.

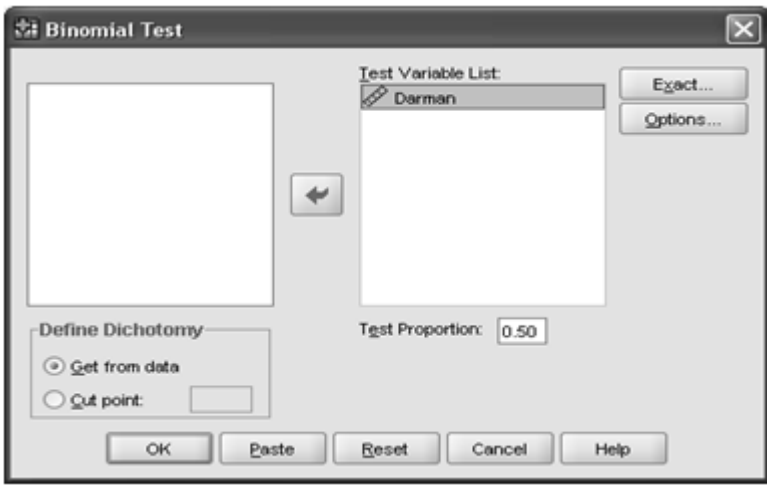
تمرین عملی

در مثالی که در ابتدای بحث مربوط به روش درمان پزشک به آن اشاره کردیم فرض کنید پزشک 10بیمار با مورد معالجه قرار داده که 7نفر آنها بهبود پیدا کرده اند و 3نفر دیگر بهبود پیدا نکرده اند. حال اگر عدد 1را به بهبود یافتگان و عدد 2را به عدم بهبود اختصاص دهید می توانید داده های این آزمایش را به spssوارد کنید و مراحل کار را به صورت زیر دنبال نمایید.
-با فرمان File/New/Dataیک فایل جدید داده باز کنید و یک متغیر دو حالتی 1) برای بهبود یافته و 2برای بهبود نیافته) با نام darmanتعریف کنید.
-در برگه Data Viewبرای متغیر Darman، به صورت زیر 7بار عدد 1و 3با عدد 2را وارد کنید.



-از مسیر زیر به کادر محاوره آزمون دو جمله‌ای وارد شوید.

Analyze/ Nonparametric Test/ Binomial Test



-متغیر darmanرا به فهرست Test Variable Listمنتقل کنید.
-در قسمت Test Protectionنسبت آزمون را 5/0وارد کنید.
-OKرا کلیک کنید.
نتایج آزمون دو جمله ای مربوط به 10بیمار که توسط پزشک معالجه شده اند در جدول زیر آمده است.

آزمون دو جمله ایBinomial Test

	گروه ها Category	تعدادN	نسبت مشاهده شده Prop. Observed	نسبت درمانهای قبلی Test Prop.	معیار تصمیم Exact Sig. (2-tailed)
Darman Group 1	1	7	.70	.50	.344
Group 2	2	3	.30		
Total		10	1.00		

در این ده نمونه 7مورد با کد 1یعنی بهبود یافته و 3مورد با کد صفر یعنی درمان نشده، مشخص شده است Test Prop .همان نسبت بهبود یافتگان در درمان های قبلی است. یعنی 50%که می‌خواهیم بر علیه آن آزمون انجام دهیم.
همچنین Observe Propنسبت بهبود یافتگان درنمونه است. که پزشک به این نتیجه رسیده است. احتمال بدست آوردن نتایجی به بزرگی مقدار مشاهده شده یابزرگتر در نمونه، هنگامی که احتمال واقعی بهبودی 50%است با عنوان Exact Sig (2-Taild)نشان داده شده است. سطح معنی داری مشاهده شده به شما می‌گوید که احتمال بدست آوردن میزان بهبودی 70%یا بزرگتر و یا 30%و کمتر هنگامی که میزان بهبودی واقعی 50%درصد است، برابر 0.34است. از آنجائی که این مقدار از 0.05بیشتر است ، نشان می‌دهد که مقدار بدست آمده بهبودی توسط پزشک خیلی غیرمعمول و دور از انتظار نیست. در حقیقت بیشتر از 34%نمونه های این جامعه دارای نتیجه غیرمعمولی به بزرگی 70%دراین نمونه می‌باشند. پس نمی‌توان فرضیه‌ای که می‌گوید روش پزشک مناسب تر از روش قبلی است، پذیرفت. اختلافی که مشاهده می‌کنید ناشی از شانس و تصادف

است. (ممکن است تغییرات مشاهده شده مربوط به نمونه گیری باشد).
از این آزمون بیشتر در مواقعی استفاده می شود که ادعایی در مورد یک نسبت صورت می گیرد و می خواهیم درستی چنین ادعایی را آزمون کنیم.
یک نکته اساسی در این آزمون که باید به آن خیلی توجه داشت، تعداد نمونه است که بر خلاف مثال بالا باید بیشتر از 25باشد زیرا توان آزمون با تعداد نمونه نسبت مستقیم داشته و به شدت به آن وابسته است.

پیشنهادات ویژه

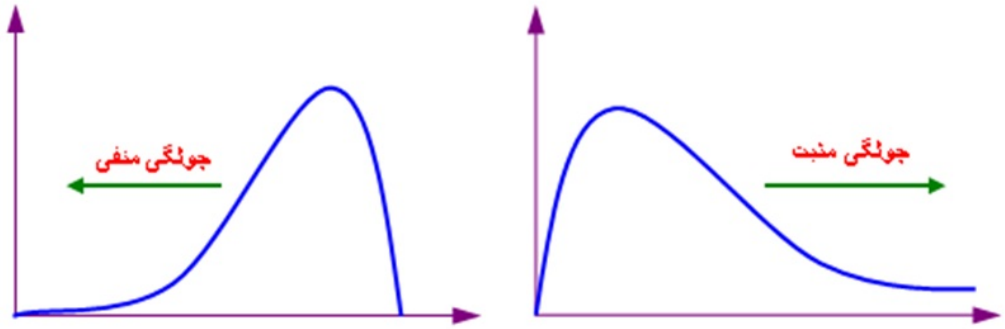
آزمون نرمال بودن داده ها

منظور از نرمال بودن توزیع داده ها این است که هیستوگرام فراوانی داده ها تقریبا به صورت منحنی نرمال باشد. پیش فرض های آماری، پایه بسیاری از آزمون های آماری تک متغیری و چندمتغیری است. شرایط مهم و اساسی برای تحلیل داده های چندمتغیری، برقراری پیش فرض های نرمال بودن، خطّی بودن و یکسانی پراکندگی داده هاست. چنانچه یک یا چندتا از این مفروضه ها نادیده گرفته شود، در این صورت، در نتایج آماری سوگیری یا تحریف رخ می دهد. قبل از انجام تحلیل های آماری تک متغیره و چندمتغیره که پیش فرض های نرمال بودن، خطّی بودن، یکسانی پراکندگی و نبود هم خطّی چندگانه بین متغیرهای مستقل در مورد آن ها باید صدق کند، باید برقراری پیش فرض های آماری را بیازماییم. اگر انحراف از پیش فرض های آماری ناچیز باشد می توان با کمی تسامح و تساهل این انحراف را نادیده گرفت و به ادامه تحلیل پرداخت. اگر انحراف از پیش فرض ها قابل توجه باشد باید یا از روش تبدیل داده ها برای برقرارکردن مجدد پیش فرض ها استفاده کنیم یا از آزمون های جایگزین مانند آزمون های ناپارامتریک استفاده کنیم که پیش فرض های فوق را مطرح نمی کنند. در این بخش چگونگی نرمال بودن داده ها در نرم افزار spssبررسی می شود.

گام اول

در کادر محاوره ای که باز می شود متغیر هایی که می خواهید چولگی و کشیدگی آن را آزمون کنید را به کادر سفید انتقال دهید.
سپس روی کلید optionsکلیک کنید و در کادر محاوره ی آن گزینه های Skewnessو kurtosisرا فعال کنید.

چولگی برابر با گشتاور سوم نرمال شده است. چولگی در حقیقت معیاری از وجود یا عدم تقارن تابع توزیع می باشد. برای یک توزیع کاملاً متقارن چولگی صفر و برای یک توزیع نا متقارن با کشیدگی به سمت مقادیر بالاتر چولگی مثبت و برای توزیع نا متقارن با کشیدگی به سمت مقادیر کوچکتر مقدار چولگی منفی است. در شکل زیر چولگی مثبت و منفی را می بینید.



کشیدگی یا کورتزیس نشان دهنده قله‌مندی یک توزیع است. مقدار کشیدگی را با گشتاور چهارم نرمال بر آورد کرده اند، به عبارت دیگر کشیدگی معیاری از تیزی منحنی در نقطه ماکزیمم است و مقدار کشیدگی برای توزیع نرمال برابر ۳ می باشد. کشیدگی مثبت یعنی قله ی توزیع مورد نظر از توزیع نرمال بالاتر و کشیدگی منفی نشانه ی پایین تر بودن قله از توزیع نرمال است. در حالت کلی معمولا چنان چه چولگی و کشیدگی در بازه ی (۲ ، ۲-) نباشند داده ها از توزیع نرمال بسیار دور بوده و می بایست قبل از هر گونه آزمونی که برای انجامشان باید فرض نرمال بودن داده ها برقرار باشند؛ اصلاح گردند. (البته ممکن است بعضی از آمار دادنان این بازه را کوچکتر یا بزرگتر در نظر بگیرند.(

مثال زیر را در نظر بگیرید. از سری داده های آماده ی SPSSفایل adl.savرا از مسیر زیر باز کنید<data <Open :
در کادر open dataدر look inبه آدرس زیر رفته و فایل adl.savرا انتخاب کنید:

C:\ > program files> SPSSInc > SPSS> Samples
را باز کنید و این دو متغیر را به کادر سفید انتقال دهید و در منوی optionsدو گزینه Kurtosisو Skewnessرا فعال کنید. و در نهایت دکمه ی okرا بزنید. با این کار خروجی زیر را دریافت می کنید:

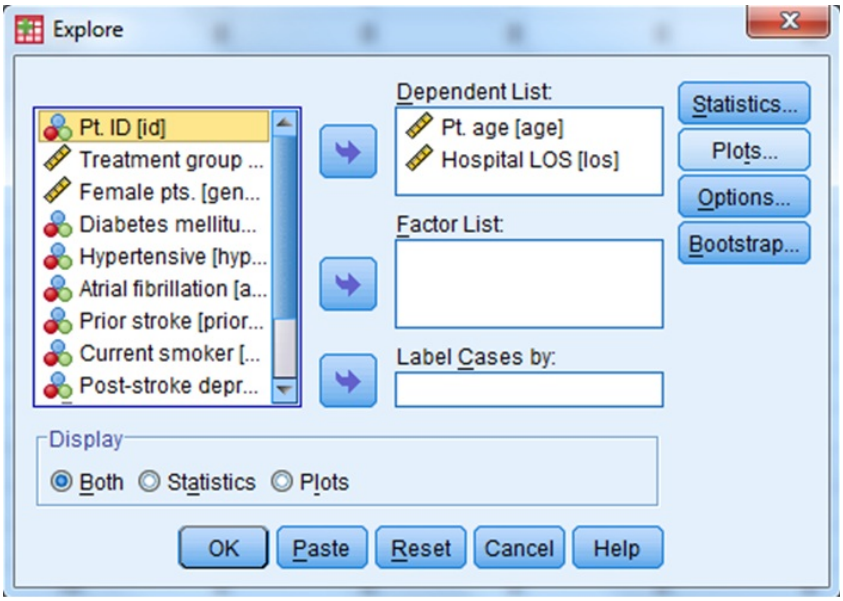
Descriptive Statistics

	N	Minimum	Maximum	Mean	Std. Deviation	Skewness		Kurtosis	
	Statistic	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error	Statistic	Std. Error
Pt. age	100	66	91	71.76	3.944	1.483	.241	4.779	.478
Hospital LOS	100	12	25	17.25	2.595	.094	.241	-.192	.478
Valid N (listwise)	100								

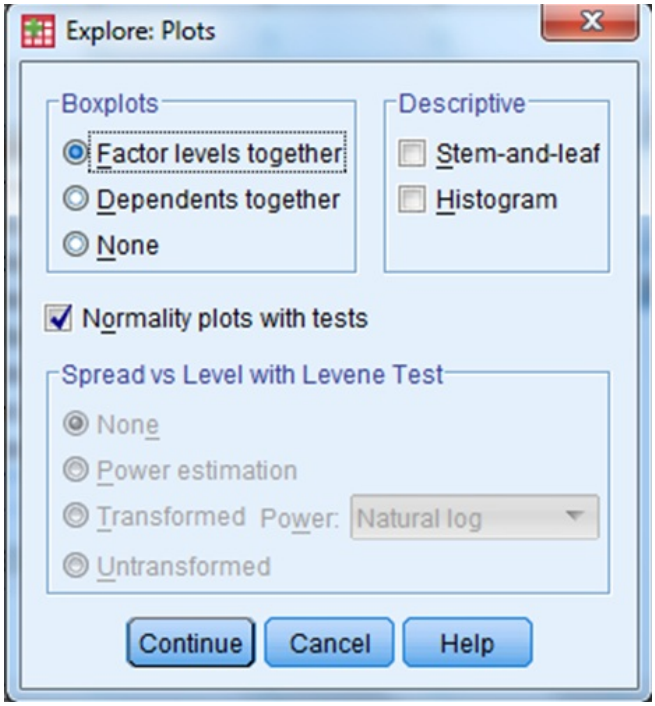
مقدار چولگی مشاهده شده برای متغیر سن 1.483است این می تواند ما را به نرمال بودن توزیع این متغیر امیدوار کند یعنی از لحاظ کجی متغیر سن مانند نرمال بوده و توزیع آن متقارن است اما مقدار کشیدگی آن از 3بیشتر است و این می رساند که قله ی این توزیع از نرمال بالاتر قرار می گیرد. لذا این متغیر با اینکه دارای توزیع متقارنی است اما نرمال نخواهد بود.
اما مقدار چولگی و کشیدگی برای متغیر losدر بازه ی (-2 , 2)قرار داشته و می توان گفت که این متغیر می تواند نرمال باشد.

گام دوم

پس از بررسی عادی یا نرمال بودن کشیدگی و یا چولگی توزیع داده هایتان، به سراغ آزمون شاپیرو ویلک بروید تا از نرمال بودن داده هایتان مطمئن گردید. برای این کار از مسیر زیر وارد کادر محاوره ی زیر شود.



در مثال بالا همان دو متغیر سن age و los را مطابق شکل وارد لیست متغیر های وابسته کنید و سایر جاها را خالی بگذارید.
سپس به منوی plots رفته و گزینه ی Normality plots with tests را تیک دار کنید.



با این عمل خروجی شما شامل جدولی است تحت عنوان Tests of Normality که به شما دو مقدار سطح معناداری را برای هر کدام از متغیر ها به طور مجزا می دهد. این مقادیر در تشخیص نرمالیتی داده ها بسیار تعیین کننده است.

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Pt. age	.125	100	.001	.891	100	.000
Hospital LOS	.122	100	.001	.975	100	.054

a. Lilliefors Significance Correction

معمولا چنانچه سطح معناداری در آزمون Shapiro-Wilk که در این جدول با .sig نمایش داده می شود بیشتر از 0.05 باشد می توان داده ها را با اطمینان بالایی نرمال فرض کرد. در غیر این صورت نمی توان گفت که داده ها توزیعشان نرمال است. با توجه به جدول فوق و مقادیر سطح معناداری برای متغیرهای age و los می توان گفت که توزیع متغیر los می تواند با احتمال خوبی نرمال باشد اما همان طوری که در گام اول هم پیش بینی کرده بودیم متغیر age نرمال نخواهد بود.

بد نیست نگاهی هم به سطح معناداری بخش Kolmogotov-Smirnov داشته باشیم هرچند این آزمون بیشتر برای مجموعه داده هایی با حجم بالا کاربرد دارد.

آزمون استقلال خی-دو

در ابتدا تعریف جدول توافقی و در ادامه آزمون استقلال مبتنی بر آزمون خی دو را شرح می دهیم:

جدول فراوانی توام دو متغیر کیفی که نشان دهنده ی وابستگی دو متغیر می باشد، جدول توافقی نامیده می شود. ما با آزمون استقلال به دنبال بررسی استقلال یا وابستگی متغیرهای کیفی در سطر و ستون ها هستیم. این آزمون به آزمون خی دو یا کای دو معروف می باشد. آماره آزمون با تکیه بر فراوانی های مشاهده شده و فراوانی ها مورد انتظار تشکیل می گردد. این آزمون همانند اکثر آزمون های آماری نیاز به پیش فرض هایی دارد که یکی از آنها اینست که فراوانی مورد انتظار در هر خانه جدول کمتر از 5نباشد و در صورت وقوع چنین موردی معمولا ادغام سطرها یا ستون هایی که مقادیر کمتر از 5دارند روشی معمول است. همانطور که در تایید این آزمون در SPSSمد نظر قرار می گیرد، چنانچه فراوانی های مورد انتظار بیش از 20درصد خانه های جدول کمتر از 5باشد مقدار معنی داری آزمون درست نخواهد بود. علاوه بر تمامی اینها می توان در جداول دو در دو از آزمون دقیق فیشر استفاده نمود که میزان معنی داری آن در خروجی نرم افزار SPSSنیز نمایش داده می شود. چنانچه هیچ یک از این روشها را نتوان بکار برد استفاده از مدل های آماری دیگر برابر بررسی رابطه استفاده می گردد.

برای انجام این آزمون در SPSSمسیر زیر را طی می کنیم:

Analyze---> Descriptive Statistics--> Crosstabs

و نهایتا جعبه گفتگوای باز می گردد که می بایست متغیرهای کیفی مورد نیاز را به قسمت های Rowو Columnفراخوانی نماییم. به طور معمول و رویه ای استاندارد در تحلیل داده ها، با انتخاب گزینه Cellsحداقل یکی از گزینه های بخش Percentagesبرای محاسبه درصد سطی، ستونی یا درصد از کل فراوانی مربوط به خانه های جداول توافقی مورد توجه قرار می گیرد. حال پاسخ سوال اصلی: برای انجام آزمون استقلال لازم است که گزینه Statisticsو سپس گزینه Chi-Squareرا انتخاب نمود.

حال در خروجی جدول آزمون کای دو، ردیف اول آماره کای دو را به عنوان کای دو پیرسون خواهیم داشت و فرض استقلال را با مشاهده مقدار احتمال (p-value)می توان بررسی نمود. چنانچه سطح معنی داری در طبیعت مساله مورد تحقیق 5درصد باشد و مقدار احتمال کمتر از 5درصد باشد فرض صفر را رد می کنیم. رد فرض صفر در آزمون استقلال به معنی اینست که دو متغیر کیفی با یکدیگر در ارتباط هستند.
در انتها ذکر می نماییم که:
-۱تنها برای جداول 2در 2آماره دقیق فیشر محاسبه نمی شود بلکه در تمامی جداول توافقی با متغیرهای کیفی بیش از 2سطح نیز این آزمون قابل محاسبه است و SPSSآن را در خروجی به محققین ارائه می دهد. در زمان گذشته به دلیل محاسبات سنگین تنها در همان جداول 2در 2آزمون دقیق مورد بحث قرار می گرفت.
-۲می توان جداول توافقی را در چند لایه، یعنی بیش از دو متغیر کیفی بررسی نمود.
-۳می توان به گسسته سازی متغیرهای کمی، آزمون استقلال را برای آن ها مورد توجه قرار داد. مثلا نمره کلاسی را به ترتیب از بیشترین نمره به کمترین به گروه های A، B، Cو Dتقسیم نمود.