معاونت تحقیقات و فناوری دانشگاه علوم پزشکی گیلان

واحد توسعه تحقیقات بالینی پورسینا

کارگاه " آموزش مقدماتی نرم افزار SPSS "

مدرس : سرکار خانم دکتر رقیه زارع

دکترای آمار زیستی

تاریخ :

۱۳۹۹/۱۱/۲۱

بخش چهارم– آزمون های آماری

ذکر نرم افزار، روش توصیف داده ها، آزمون های آماری و محاسبات اصلی مورد استفاده برای آنالیز در این قسمت مهم است.

Type of Error

Decision	True state	
	H_0 is True	H_1 is True
Reject H_0	Type I error	No error ☺
Fail to reject H_0	No error ☺	Type II error

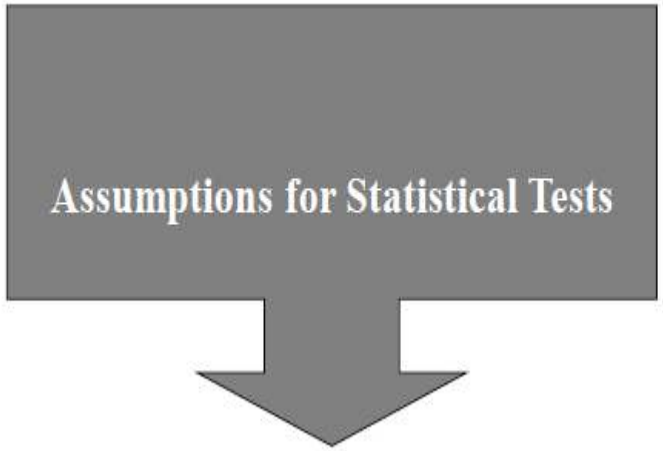
The probability of a type I and II errors are respectively labeled α and β.

Alpha	Zα (One tailed)	Zα/2 (two tailed)
0.20	0.84	1.28
0.15	1.04	1.44
0.10	1.28	1.64
0.05	1.64	1.96
0.01	2.33	2.58
0.001	3.09	3.29

The normal deviates for Type I error (Alpha)

Power	$Z_{1-\beta}$
0.70	0.52
0.75	0.67
0.80	0.84
0.85	1.03
0.90	1.28
0.95	1.64
0.99	2.33

The normal deviates for statistical power



Unfortunately, even simple and basic procedures such as the popular t-tests or chi-square tests are frequently misused in medical research, because

their test assumptions are not evaluated sufficiently before application.

آزمونهای آماری

آزمون آماری استنباطی بر اساس مقیاس اندازه گیری متغیرها به دو نوع پارامتری و ناپارامتری می باشند. این دو نوع مقیاس دارای تفاوت هایی با یکدیگر می باشند که در جدول زیر به آنها اشاره شده است:

نام آزمون	آزمون‌های پارامتری	آزمون‌های ناپارامتری
نوع اطلاعاتی که بررسی می کند	مقیاس فاصله‌ای و نسبی	مقیاس اسمی و رتبه‌ای
شاخص آماری	میانگین (Mean) و واریانس (Variance)	میانه (Median) و نما (Mode)
نوع پیش فرض مورد نیاز	توزیع جامعه نرمال	مستلزم هیچگونه فرضی در مورد توزیع نیست. *

خلاصه آزمونهای پارامتریک

نوع آزمون	کاربرد
جامعه t تک نمونه	زمون فرض پیرامون میانگین یک جامعه بررسی فرضیه های پژوهش و تحلیل سوالات تخصصی آن
جامعه t وابسته	پیرامون دو میانگین از یک جامعه
جامعه t دو نمونه مستقل	جهت مقایسه میانگین دو جامعه (واریانس دو جامعه برابر نیست)
جامعه t ولچ	جهت مقایسه میانگین دو جامعه (واریانس دو جامعه برابر نیست)
جامعه t هتلینگ	مقایسه چند میانگین از دو جامعه
تحلیل واریانس (ANOVA)	بررسی اختلاف میانگین چند جامعه آماری
تحلیل واریانس چندعاملی (MANOVA)	بررسی اختلاف چند میانگین از چند جامعه آماری
تحلیل کوواریانس چندعاملی (MANCOVA)	بخواهیم اثر یک یا چند متغیر کمکی را حذف کنیم

خلاصه آزمونهای پارامتریک

- آزمون t تک نمونه : برای آزمون فرض پیرامون میانگین یک جامعه استفاده می شود. در بیشتر پژوهش هائی که با مقیاس لیکرت انجام می شوند جهت بررسی فرضیه های پژوهش و تحلیل سوالات تخصصی مربوط به آنها از این آزمون استفاده می شود.
- آزمون t وابسته : برای آزمون فرض پیرامون دو میانگین از یک جامعه استفاده می شود. برای مثال اختلاف میانگین رضایت کارکنان یک سازمان قبل و بعد از تغییر مدیریت یا زمانی که نمرات یک کلاس با پیش آزمون و پس آزمون سنجش می شود.
- آزمون t دو نمونه مستقل: جهت مقایسه میانگین دو جامعه استفاده می شود. در آزمون t برای دو نمونه مستقل فرض می شود واریانس دو جامعه برابر است. برای نمونه به منظور بررسی معنی دار بودن تفاوت میانگین نمره نظرات پاسخ دهندگان بر اساس جنسیت در خصوص هر یک از فرضیه های پژوهش استفاده میشود.
- آزمون t ولچ: این آزمون نیز مانند آزمون t دو نمونه جهت مقایسه میانگین دو جامعه استفاده می شود. در آزمون t ولچ فرض می شود واریانس دو جامعه برابر نیست. برای نمونه به منظور بررسی معنی دار بودن تفاوت میانگین نمره نظرات پاسخ دهندگان بر اساس جنسیت در خصوص هر یک از فرضیه های پژوهش استفاده میشود.

خلاصه آزمونهای پارامتریک

- آزمون tهتلینگ : برای مقایسه چند میانگین از دو جامعه استفاده می شود. یعنی دو جامعه براساس میانگین چندین صفت مقایسه شوند.
- تحلیل واریانس (ANOVA) از این آزمون به منظور بررسی اختلاف میانگین چند جامعه آماری استفاده می شود. برای نمونه جهت بررسی معنی دار بودن تفاوت میانگین نمره نظرات پاسخ دهندگان بر اساس سن یا تحصیلات در خصوص هر یک از فرضیه های پژوهش استفاده می شود.
- تحلیل واریانس چندعاملی (MANOVA) از این آزمون به منظور بررسی اختلاف چند میانگین از چند جامعه آماری استفاده می شود.
- تحلیل کوواریانس چندعاملی (MANCOVA) چنانچه در MANOVAبخواهیم اثر یک یا چند متغیر کمکی را حذف کنیم استفاده می شود.
- ضریب همبستگی گشتاوری پیرسون: برای محاسبه همبستگی دو مجموعه داده استفاده می شود.
- خلاصه آزمونهای آماری ناپارامتریک

خلاصه آزمونهای نان پارامتریک

نوع آزمون	کاربرد
آزمون علامت تک نمونه	آزمون فرض پیرامون میانگین یک جامعه
آزمون علامت زوجی	آزمون فرض پیرامون دو میانگین از یک جامعه
ویلکاکسون	اختلاف نسبی تفاوت از میانگین
من-ویتنی	مقایسه میانگین دو جامعه
کروسکال-والیس	بررسی اختلاف میانگین چند جامعه آماری
فریدمن	آنالیز واریانس دو عاملی است که در آن k تیمار به صورت تصادفی به n بلوک تخصیص داده شده اند.
کولموگروف-اسمیرنف	نوعی آزمون نیکوئی برازش برای مقایسه یک توزیع نظری با توزیع مشاهده شده
آزمون تقارن توزیع	شکل توزیع مورد سوال قرار می گیرد.(توزیع متقارن نیست)
آزمون میانه	مقایسه میانه دو جامعه
مک نمار	بررسی مشاهدات زوجی درباره متغیرهای دو ارزشی
آزمون Q کوکران	تعمیم آزمون مک نمار در k نمونه وابسته
ضریب همبستگی اسپیرمن:	حاسبه همبستگی دو مجموعه داده که به صورت ترتیبی قرار دارند

خلاصه آزمونهای نان پارامتریک

- آزمون علامت تک نمونه : برای آزمون فرض پیرامون میانگین یک جامعه استفاده می شود.
- آزمون علامت زوجی : برای آزمون فرض پیرامون دو میانگین از یک جامعه استفاده می شود.
- ویلکاکسون : همان آزمون علامت زوجی است که در آن اختلاف نسبی تفاوت از میانگین لحاظ می شود.
- مان-ویتنی: به آزمون **U** نیز موسوم است و جهت مقایسه میانگین دو جامعه استفاده می شود.
- کروسکال-والیس: از این آزمون به منظور بررسی اختلاف میانگین چند جامعه آماری استفاده می شود. به آزمون **H** نیز موسوم است و تعمیم آزمون **U** مان-ویتنی می باشد. آزمون کروسکال-والیس معادل روش پارامتریک آنالیز واریانس تک عاملی است.
- فریدمن: این آزمون معادل روش پارامتریک آنالیز واریانس دو عاملی است که در آن **k** تیمار به صورت تصادفی به **n** بلوک تخصیص داده شده اند.
- آزمون خی-دو : χ^2 این آزمون کاربردهای متعددی دارد برای مثال معادل ناپارامتریک آزمون تی-تک نمونه، آزمون استقلال و نیکویی برازش
- کولموگروف-اسمیرنف : نوعی آزمون نیکوئی برازش برای مقایسه یک توزیع نظری با توزیع مشاهده شده است.

خلاصه آزمونهای نان پارامتریک

- آزمون تقارن توزیع : در این آزمون شکل توزیع مورد سوال قرار می گیرد. فرض بدیل آن است که توزیع متقارن نیست.
- آزمون میانه : جهت مقایسه میانه دو جامعه استفاده می شود و برای k جامعه نیز قابل تعمیم است.
- مک نمار : برای بررسی مشاهدات زوجی درباره متغیرهای دو ارزشی استفاده می شود.
- آزمون Qکوکران: تعمیم آزمون مک نمار در k نمونه وابسته است.
- ضریب همبستگی اسپیرمن: برای محاسبه همبستگی دو مجموعه داده که به صورت ترتیبی قرار دارند استفاده می شود.

برای بررسی داده‌های کمی پژوهش، استفاده از آزمون‌های آماری الزامی است. برای هر پژوهشی لزوماً از یک روش آماری استفاده می‌شود. مرور این آزمون‌ها می‌تواند راهنمای مناسبی برای تجزیه و تحلیل داده‌های شما باشد. به خاطر داشته باشید، ممکن است همواره استفاده از یک آزمون کافی نباشد. همچنین به خاطر داشته باشید آزمون‌های متفاوت‌تری از آن‌چیز که معرفی شده است، به کار رود. بسیاری از این آزمون‌ها امروزه با نرم‌افزارهای آماری همچون SPSSقابل انجام است.

آزمون نرمالیته داده ها

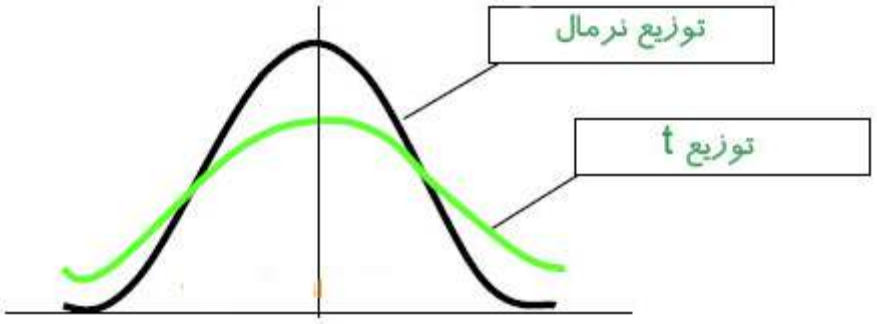
آزمون نرمال بودن داده‌ها روشی برای تشخیص آن است که مشخص شود توزیع داده‌های گردآوری شده از توزیع طبیعی یا نرمال برخوردار است. قبل از هر گونه آزمونی که با فرض نرمال بودن داده‌ها صورت می‌گیرد باید آزمون نرمال بودن صورت گیرد. برای این منظور روش‌های متعددی وجود دارد.

در آمار استنباطی شرط اصلی برای انواع آزمون‌های آمار پارامتریک و ناپارامتریک به توزیع داده‌ها بستگی دارد. اگر توزیع داده‌ها نرمال باشد در اینصورت از روش‌های پارامتریک استفاده می‌شود و اگر نرمال نباشد نباید از روش‌های پارامتریک استفاده شود. آزمون‌های ناپارامتریک ربطی به توزیع داده‌ها ندارد. برخی از آزمون های بررسی نرمال بودن داده ها عبارتند از:

- چولگی و کشیدگی داده‌ها
- ترسیم نمودار هیستوگرام
- آزمون کولموگروف-اسمیرنوف
- آزمون شاپیرو-ویلک

کاربرد چولگی و کشیدگی در آزمون نرمال بودن داده‌ها

بهترین روش برای داده‌های طیف لیکرت و پرسشنامه بررسی چولگی و کشیدگی داده ها است. چولگی معیاری از تقارن یا عدم تقارن تابع توزیع می‌باشد. برای یک توزیع کاملاً متقارن چولگی صفر و برای یک توزیع نامتقارن با کشیدگی به سمت مقادیر بالاتر چولگی مثبت و برای توزیع نامتقارن با کشیدگی به سمت مقادیر کوچکتر مقدار چولگی منفی است. کشیدگی یا kurtosis نشان دهنده ارتفاع یک توزیع است. به عبارت دیگر کشیدگی معیاری از بلندی منحنی در نقطه ماکزیمم است و مقدار کشیدگی برای توزیع نرمال برابر ۳ می باشد. کشیدگی مثبت یعنی قله توزیع مورد نظر از توزیع نرمال بالاتر و کشیدگی منفی نشانه پایین تر بودن قله از توزیع نرمال است. برای مثال در توزیع t که پراکندگی داده ها بیشتر از توزیع نرمال است، ارتفاع منحنی کوتاه تر از منحنی نر



کاربرد چولگی و کشیدگی در آزمون نرمال بودن داده‌ها

در حالت کلی چنانچه چولگی و کشیدگی در بازه (۲ ، ۲-) نباشند داده‌ها از توزیع نرمال برخوردار نیستند.

- رسم نمودار هیستوگرام برای آزمون نرمال بودن داده‌ها

ترسیم نمودار هیستوگرام از روش‌های آزمون نرمال بودن داده ها است. با استفاده از نرم افزار SPSS به سادگی می‌توان نمودار هیستوگرام با نمایش منحنی نرمال را ترسیم کرد.

- آزمون کولموگروف-اسمیرنوف

علاوه بر بررسی عادی یا نرمال بودن کشیدگی و یا چولگی توزیع داده‌ها، از آزمون شاپیرو-ویلک یا آزمون کولموگروف-اسمیرنوف استفاده می شود برای آزمون نرمال بودن داده ها استفاده می‌شود.

هنگام بررسی نرمال بودن داده‌ها ما فرض صفر مبتنی بر اینکه توزیع داده‌ها نرمال است را در سطح خطای ۵٪ تست می‌کنیم. بنابراین اگر آماره آزمون بزرگتر مساوی ۰،۰۵ بدست آید، در این صورت دلیلی برای رد فرض صفر مبتنی بر اینکه داده نرمال است، وجود نخواهد داشت. به عبارت دیگر توزیع داده‌ها نرمال خواهد بود. برای آزمون نرمالیته فرض‌های آماری به صورت زیر تنظیم می‌شود:

H0: توزیع داده‌های مربوط به هر یک از متغیرها نرمال است

H1: توزیع داده‌های مربوط به هر یک از متغیرها نرمال نیست

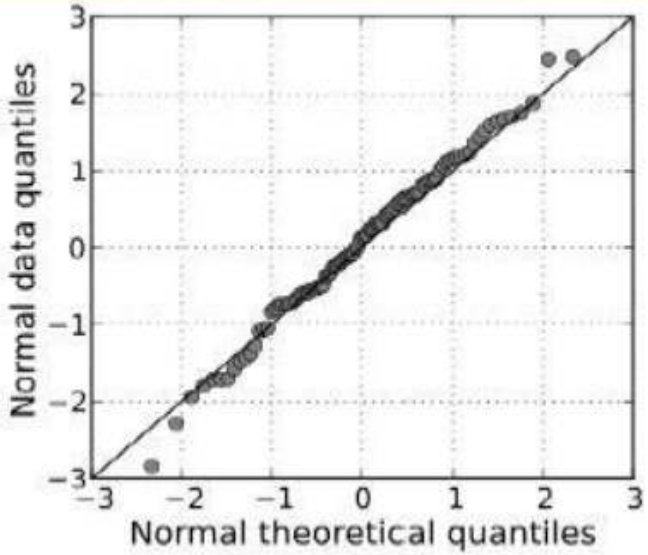
Independent T-test

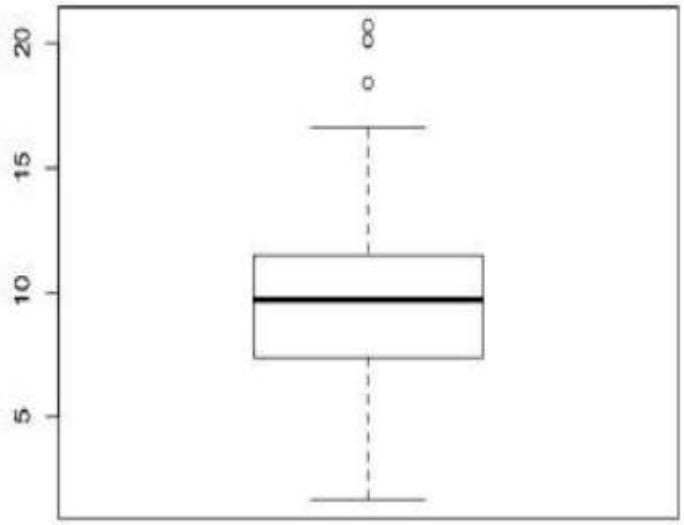
Independent T-Test: Some of the most important assumptions

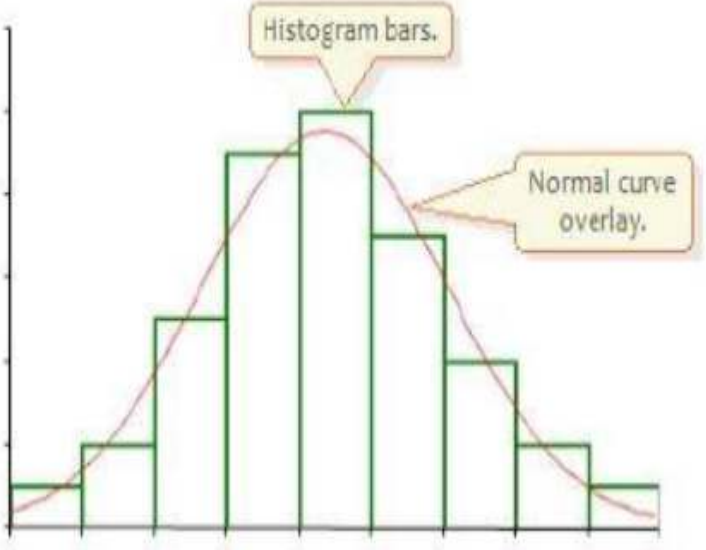
- **dependent variable: continuous scale**
- **independent variable: two categorical, independent groups.**
- **dependent variable: should be approximately normally distributed for each group of the independent variable.**
- There needs to be **homogeneity of variances. Levene's test in SPSS .**

Normality Assessment

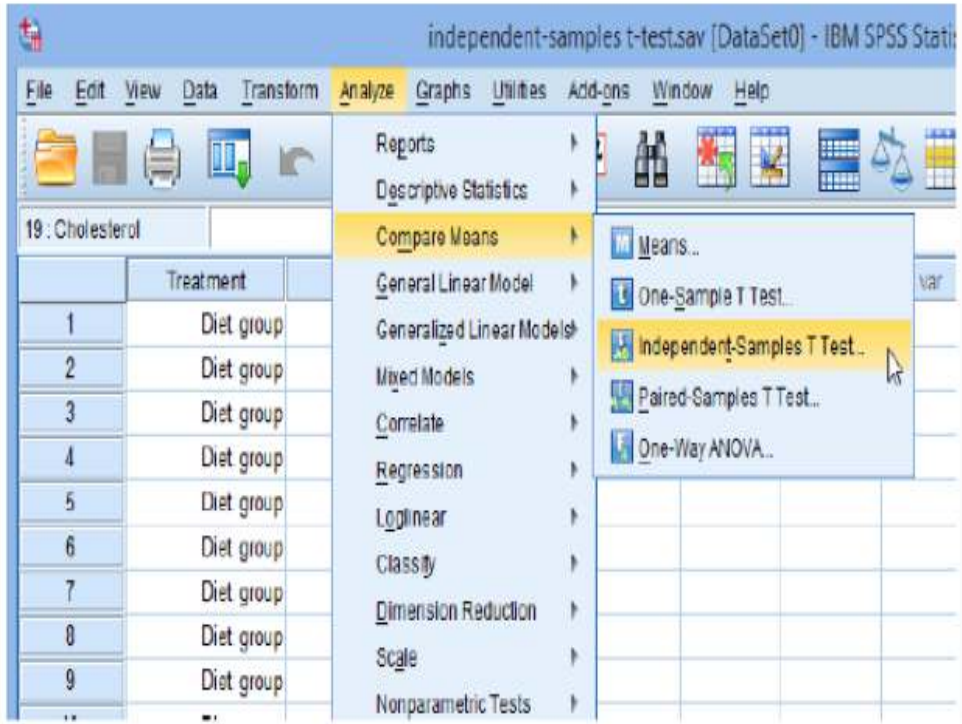
1. Graphical methods (histograms, boxplots, Q-Q-plots)
2. Numerical methods (skewness and kurtosis indices)
3. Normality tests


Normal data quantiles
Normal theoretical quantiles
3
2
1
0
−1
−2
−3
−3
−2
−1
0
1
2
3


20
15
10
5


Histogram bars.
Normal curve overlay.

Click **Analyze > Compare Means > Independent-Samples T Test...**

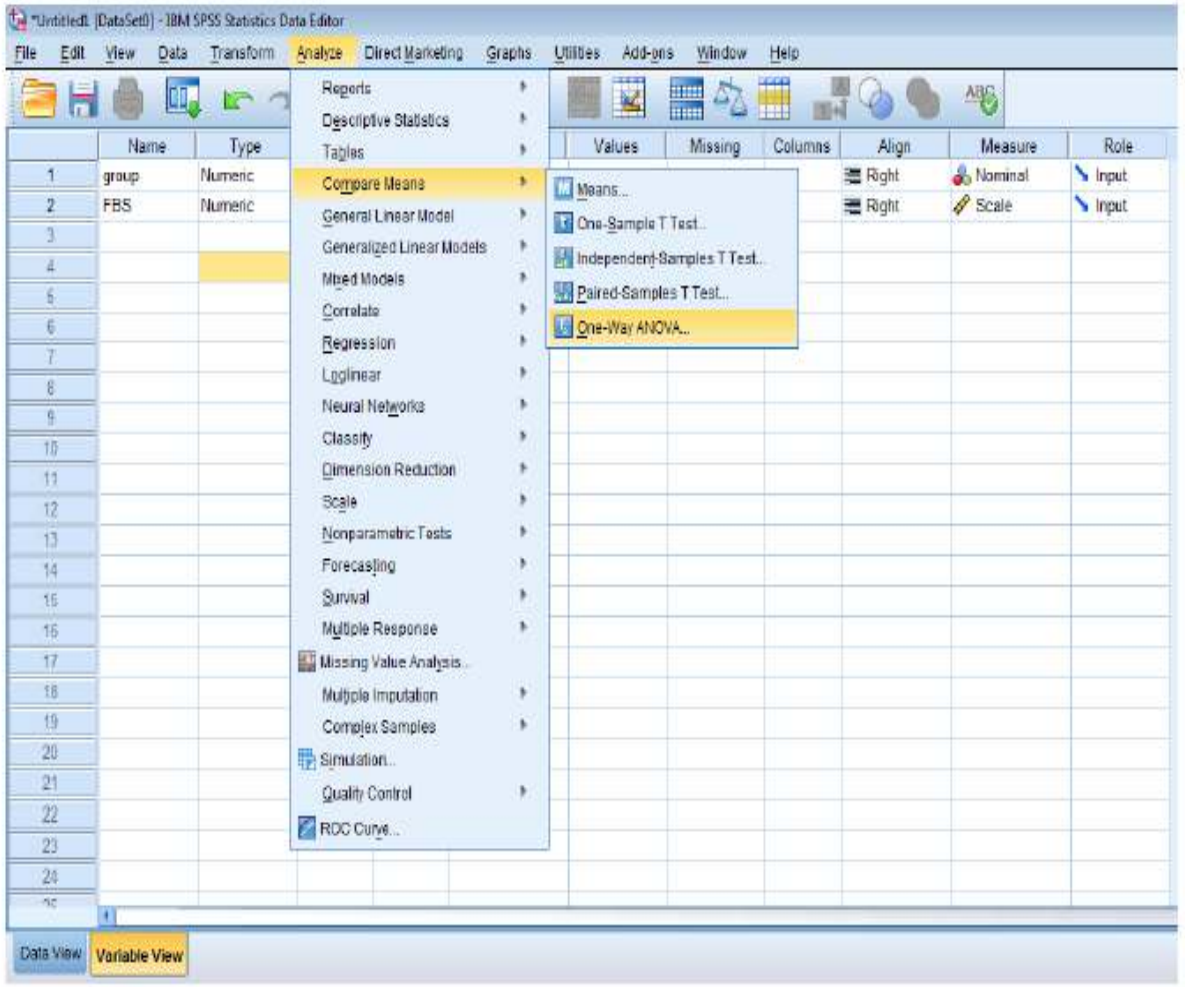


One way ANOVA

One way ANOVA: Some of the most important assumptions

- **dependent variable:** continuous scale
- **independent variable:** More than two categorical, independent groups.
- **dependent variable:** should be approximately normally distributed for each group of the independent variable.
- There needs to be homogeneity of variances.

Click **Analyze > Compare Means > one way ANOVA ...**

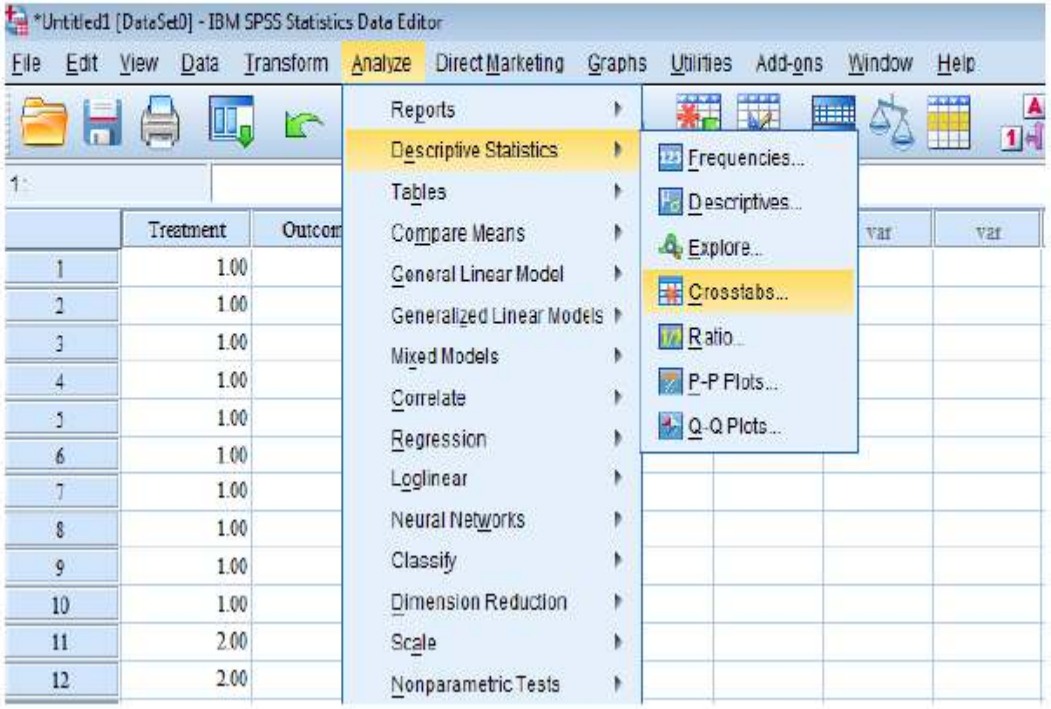


Chi-Square Test:

Chi-Square Test: Some of the most important assumptions

- **Independence**: This means chi-squared cannot be used to test correlated data (like matched pairs).
- **Adequate expected cell counts**: A common rule is 5 or more in all cells of a 2-by-2 table, and 5 or more in 80% of cells in larger tables, but no cells with zero expected count.

Click **Analyze** > **Descriptive Statistics** > **Crosstabs...**

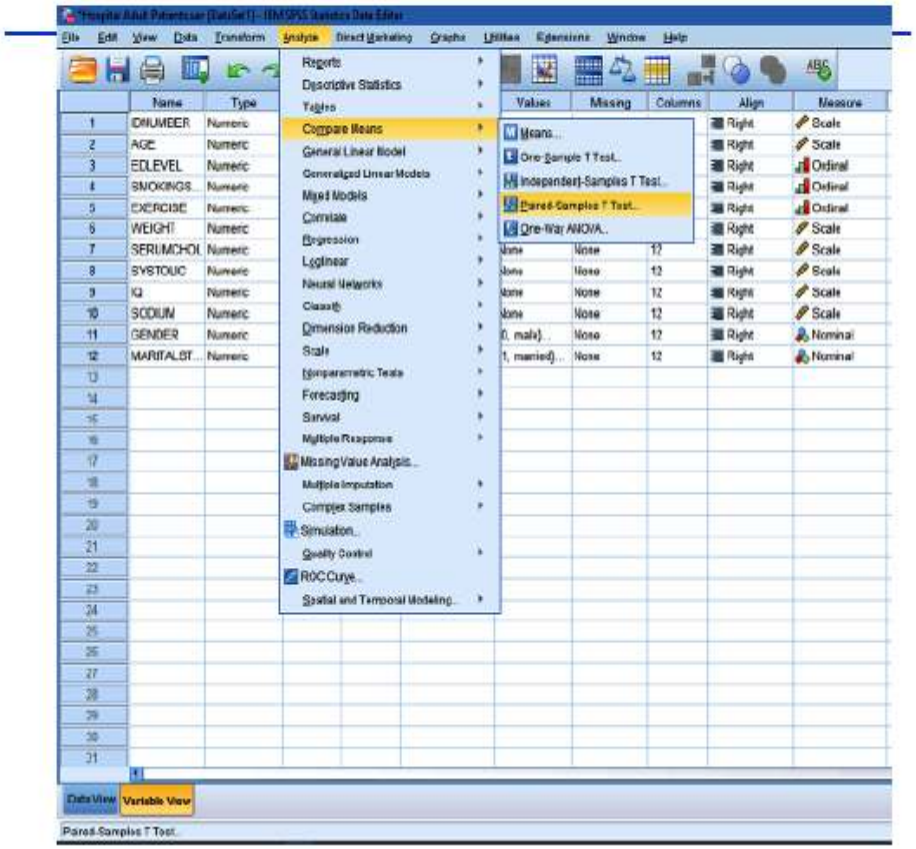


Paired t-test

Paired t-test : Some of the most important assumptions

- dependent variable: continuous scale
- independent variable: Independent variable should consist of two categorical, "related groups" or "matched pairs".
- The distribution of the differences in the dependent variable between the two related groups should be approximately normally distributed.

Click **Analyze > Compare Means > Paired-Samples T Test...**



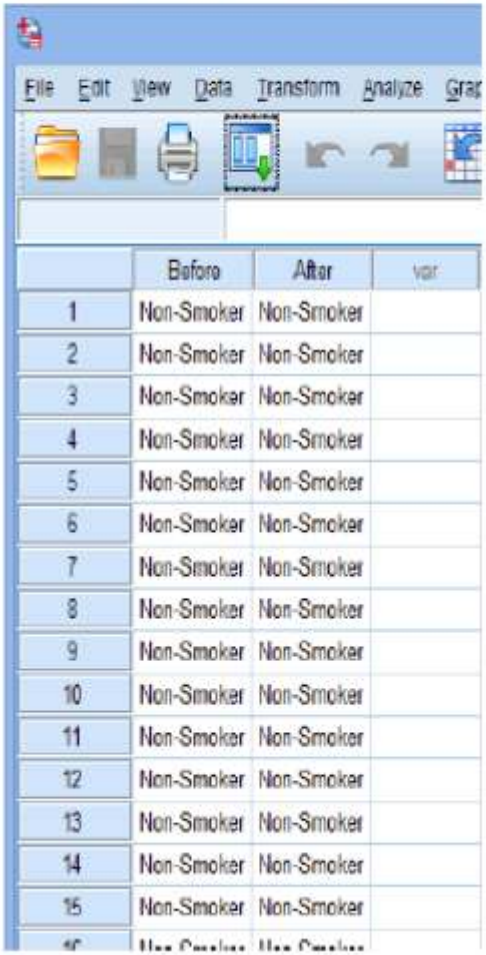
McNemar's test:

McNemar's test: Some of the most important assumptions

- You have one categorical dependent variable with two categories (i.e. A dichotomous variable) and one categorical independent variable with two related groups.
- The two groups of your dependent variable must be mutually exclusive. This means that no groups can overlap. In other words, a participant can only be in one of the two groups; they cannot be in both groups at the same time.
- The cases (e.g., participants) are a random sample from the population of interest. However, in practice, this is not always how sampling took place.

McNemar's test using SPSS

Individual scores for each participant

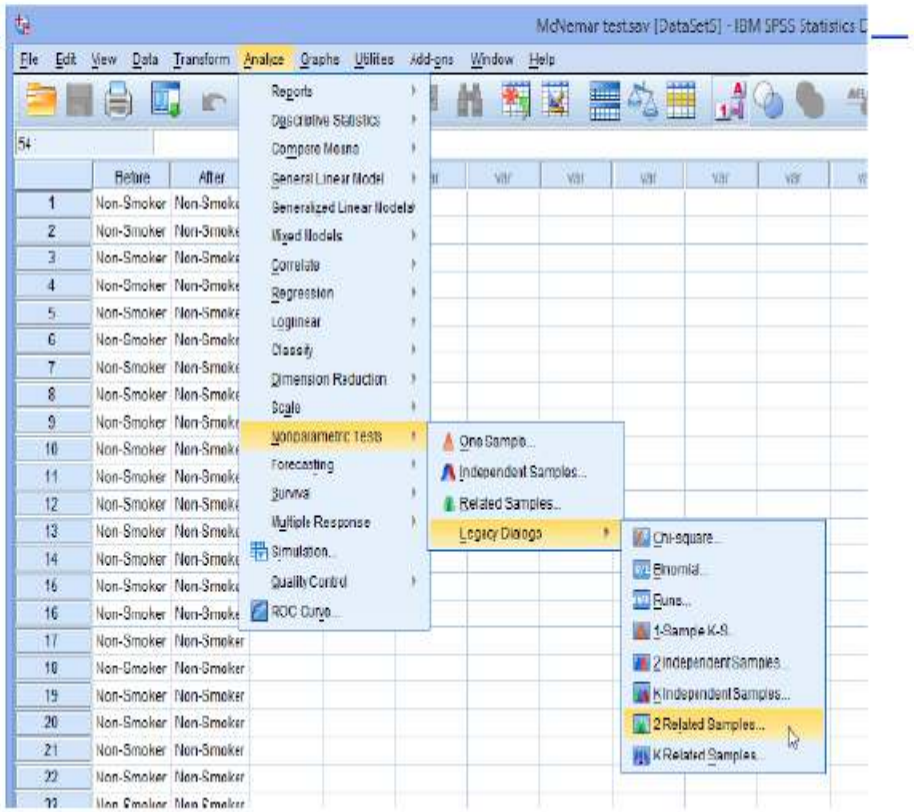


	Before	After	var
1	Non-Smoker	Non-Smoker	
2	Non-Smoker	Non-Smoker	
3	Non-Smoker	Non-Smoker	
4	Non-Smoker	Non-Smoker	
5	Non-Smoker	Non-Smoker	
6	Non-Smoker	Non-Smoker	
7	Non-Smoker	Non-Smoker	
8	Non-Smoker	Non-Smoker	
9	Non-Smoker	Non-Smoker	
10	Non-Smoker	Non-Smoker	
11	Non-Smoker	Non-Smoker	
12	Non-Smoker	Non-Smoker	
13	Non-Smoker	Non-Smoker	
14	Non-Smoker	Non-Smoker	
15	Non-Smoker	Non-Smoker	

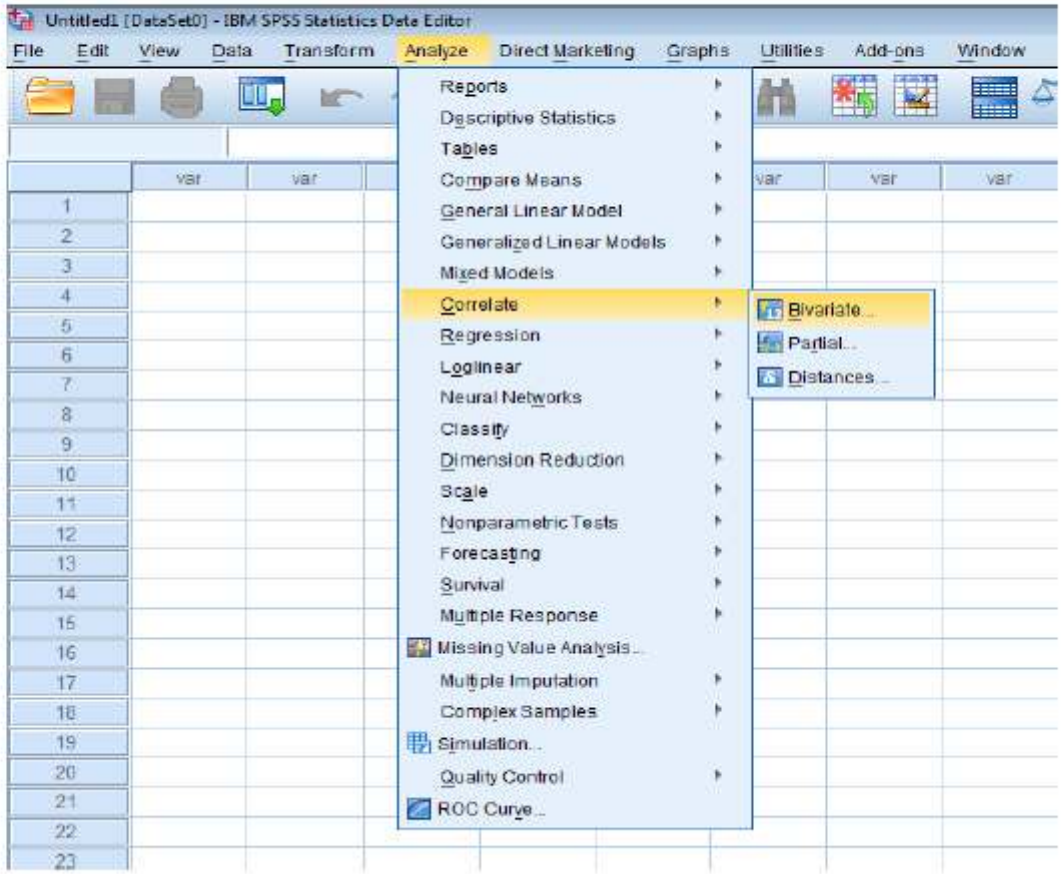
Total count data (frequencies)

	Before	After	Freq
1	Non-Smoker	Non-Smoker	20
2	Non-Smoker	Smoker	5
3	Smoker	Non-Smoker	16
4	Smoker	Smoker	9
5			
6			
7			
8			
9			

Click **Analyze > Nonparametric Tests> legacy dialogs> 2 related samples**



Click **Analyze > Correlate> Bivariate**



نام کارگاه : نرم افزار spss **تاریخ :** ۱۳۹۹،۱۱،۲۱ **واحد توسعه تحقیقات بالینی پورسینا**

ردیف	نام و نام خانوادگی	محل خدمت	شماره ملی	امضاء
۱	سپیده نعمات ثابت	واحد توسعه بالینی	۲۵۹۵۶۷۹۰۳۱	
۲	آسیه محمدخواه	واحد آموزش پزشکی	۲۷۳۹۴۵۶۰۹۳	
۳	سینا محسنی	بلبورد کنیت	۲۵۹۳۴۷۸۶۵۹	
۴	لیلا بلالی	واحد آموزش پزشکی	۲۵۹۳۳۰۴۰۴۸	
۵	فاطمه نمازی	واحد تحقیقات بالینی	۲۵۹۳۴۷۷۶۶۰	
۶	مرضیه اکبرزاده	منابع انسانی	۲۵۹۵۷۹۲۳۳۷	
۷	آذین بدمنه	دانشکده پزشکی	۵۱۸۰۰۷۴۷۱۱	