الموضوع الرابع: الاختبارات المعلمية والاختبارات اللامعلمية

1- الاختبارات المعلمية وتطبيقاتها

2- الاختبارات اللامعلمية وتطبيقاتها

3- تصحيح تطبيقات الاختبارات المعلمية والاختبارات اللامعلمية

1- الاختبارات المعلمية (Parametric Tests)

اختبار التوزيع الطبيعي كولموغوروف – سميرنوف Kolmogorov–Smirnov Test

يستخدم هذا الاختبار لمعرفة طبيعة توزيع بيانات ظاهرة معينة كونها تتبع التوزيع الطبيعي من عدمه. وهذا الاختبار ضروري في الاختبارات المعلمية، حيث يشترط أن يكون توزيع البيانات طبيعي مع ملاحظة أنه يستخدم اختبار كولموغوروف – سميرنوف لمعرفة توزيع البيانات إذا كان حجم العينة أكبر من أو يساوي 50، بينما يستخدم اختبار شبيرو–ويلك (Shapiro–Wilk) وإذا كان حجم العينة أ صغر من50.

تطبيق: تمثل البيانات التالية درجات 50 طالبا في مقرر "الأدوات الإحصائية لتحليل البيانات":

21	32	76	82	90
40	30	65	92	80
88	45	82	60	70
89	89	80	70	90
92	88	90	50	60
85	77	92	65	76
79	86	86	79	68
31	90	71	82	94
29	94	93	68	83
50	97	68	80	74

المطلوب: استخدم اختبار كولموغوروف – سميرنوف لمعرفة أن البيانات السابقة لها توزيع طبيعي أم لا مستخدما مستوى الدلالة $\alpha = 0.05$.

اختبار t للعينة الواحدة One Sample T- Test

هدفه: اختبار t للعينة الواحدة هو اختبار معلمي يستخدم لدراسة متوسط المجتمع في حالة العينات الصغيرة $(n < 30)$، والتحقق مما إذا كانت هناك فروق معنوية بين متوسط المجتمع μ والمتوسط الفرضي μ_0 .

شروط تطبيقه:

- بيانات المجتمع مستقلة عن بعضها؛
- بيانات المتغير مقاسة على الأقل في المستوى الفتري؛

- بيانات المتغير موزعة توزيعا قريبا من التوزيع الطبيعي؛

تطبيق: البيانات التالية تمثل درجات عشرين طالبا في مساق مادة الرياضيات:

65 72 68 82 45 92 87 85 90 60 48 60 68 72 79 68 73 69 78 84

المطلوب: هل يختلف تحصيل هؤلاء الطلبة في مادة الرياضيات عن الدرجة 62 عند مستوى الدلالة 5%؟

اختبار (t) لعينتين مستقلتين Independent Samples t- test

هدفه: اختبار t لعينتين مستقلتين هو اختبار معلمي يستخدم لدراسة متوسطي مجتمعين مستقلين والتحقق مما إذا كانت هناك فروق معنوية بين متوسطي المجتمعين أم لا.

شروط تطبيقه:

في حالة العينات الصغيرة $(n < 30)$، يشترط أن تكون:

- بيانات المجتمع الأول مستقلة عن بيانات المجتمع الثاني؛
- المتغير التابع مقاس على الأقل في المستوى الفتري؛
- المتغير التابع موزع توزيع توزيعا قريبا من التوزيع الطبيعي في كل مجموعة؛
- تجانس التباين بين المجموعتين.

تطبيق: رغب أستاذ في تطوير مستوى الطالب من خلال معرفة أي الطريقتين أفضل، الطريقة التقليدية التي تعتمد على 30% في المخبر، أم الطريقة الجديدة التي تعتمد على 70% في المخبر. ولغرض التوصل إلى قرار مناسب اختار عينتين A و B بشكل مستقل وبعد مدة أجرى امتحان موحد لهما فكانت النتائج على النحو التالي:

A	85	70	60	75	65	75	78	77	65	80
B	70	85	90	65	75	85	80	88	82	

هل هناك فروق بين المجموعتين A و B في نتائج الامتحان عند مستوى دلالة 0.05؟

اختبار (t) لعينتين مرتبطتين Paired Samples t- test

هدفه: اختبار t لعينتين مرتبطتين هو اختبار معلمي يستخدم لدراسة متوسطي مجتمعين مرتبطين والتحقق مما إذا كانت هناك فروق معنوية بين متوسطي المجتمعين أم لا.

شروط تطبيقه:

في حالة العينات الصغيرة $(n < 30)$، يشترط أن تكون:

- بيانات المتغير مقاسة على الأقل في المستوى الفتري؛
- الفروق بين درجات القياسين موزعة توزيعا قريبا من التوزيع الطبيعي؛

تطبيق: تم إشراك 14 موظفا في دورة تكوينية وقيست معلوماتهم قبل وبعد اشتراكهم في الدورة، وكانت النتائج كالتالي:

رقم	الدرجة قبل الدورة	الدرجة بعد الدورة	رقم	الدرجة قبل الدورة	الدرجة بعد الدورة
1	209	217	8	217	230
2	215	219	9	231	240
3	215	210	10	239	242
4	227	232	11	223	214
5	22	226	12	228	227
6	212	219	13	223	235
7	234	238	14	223	226

المطلوب: هل للدورة التكوينية تأثير على مستوياتهم المعلوماتية عند مستوى دلالة 5%؟

تحليل التباين الأحادي باتجاه واحد One-Way ANOVA

هدفه: اختبار تحليل التباين الأحادي هو اختبار معلمي هام ويدعى اختبار F، يستخدم لدراسة الفروق بين المتوسطات لثلاث عينات مستقلة أو أكثر للتحقق مما إذا كانت هناك فروق معنوية بين متوسطات العينات قيد الدراسة أم لا.

شروط تطبيقه

في حالة العينات الصغيرة $(n < 30)$، يشترط أن تكون:

- العينات مستقلة عن بعضها البعض مثنى مثنى؛
- المتغير التابع مقاس على الأقل في المستوى الفتري؛
- المتغير التابع موزع توزيع قريبا من التوزيع الطبيعي في كل مجموعة؛
- تجانس التباين بين المجموعات (تساوي التباينات).

تطبيق: استخدمت أربع طرق صناعية لإنتاج نوع معين من القماش بثلاث مكررات لكل طريقة وكانت النتائج حسب الجدول التالي:

المكررات / الطريقة	1	2	3	المتوسط
الطريقة 1	55	57	48	50
الطريقة 2	55	64	64	61
الطريقة 3	55	49	52	52
الطريقة 4	50	44	41	45

المطلوب: – هل توجد فروق جوهرية بين متوسطات الطرق الصناعية لإنتاج القماش عند مستوى دلالة 5%؟ في حالة ظهور فروق معنوية بين الطرق الصناعية، اختبر معنوية الفروق بين متوسطي كل طريقتين باستخدام طريقة الفرق المعنوي الأصغر $(L.S.D)$ لمستوى الدلالة 5%.

2- الاختبارات اللامعلمية (Nonparametric Tests)

اختبار جودة المطابقة Goodness of fit test لعينة واحدة

هدفه: اختبار جودة المطابقة هو اختبار لامعلمي، يستخدم للتأكد من أن نموذجا ما يناسب البيانات بشكل جيد. ويستخدم لدراسة التوزيع الاحتمالي للمتغير، ويعتمد على مقارنة القيم المشاهدة $(Observed)$ والقيم المتوقعة (Expected) من التوزيع الاحتمالي.

شرط تطبيقه: أن تكون جميع القيم المتوقعة أكبر من 5.

ويعطي SPSS تحت جدول المخرجات عدد الخلايا المتوقعة ذات القيم الأقل من 5 ونسبتها المئوية، حيث يمكن أحيانا قبول نسبة 20% من القيم المتوقعة الأقل من 5.

تطبيق: تمثل البيانات التالية عدد الأشخاص الذين تناولوا طعام العشاء في مطعم صغير على مدى 50 يومًا:

25	7	10	8	16	24	22	8	12	10	5	14	27
15	20	12	16	19	24	6	10	1	15	23	8	30
19	16	8	6	9	7	12	14	19	22	20	16	14
		20	21	16	18	12	16	23	20	4	17	27

الموضوع الرابع: الاختبارات المعلمية والاختبارات اللامعلمية

المطلوب: هل متغير عدد الأشخاص الذين تناولوا العشاء في المطعم يتبع التوزيع الطبيعي عند مستوى الدلالة 5%؟

اختبار كاي تربيع Chi-Square test للاستقلالية

نقوم في كثير من المسائل العملية، بتصنيف مجموعة من المشاهدات وفق أسلوبين، فينشأ السؤال التالي: هل هناك علاقة بين أسلوبي التصنيف؟ مثل:

- ✓ هل هناك علاقة بين الجنس والمستوى الأكاديمي؟ و وهل هناك علاقة بين التدخين والإصابة بسرطان الرئة؟

للإجابة عن هذه الأسئلة وأمثالها نستعمل اختبار كاي تربيع للاستقلالية.

تطبيق: في دراسة للعلاقة بين التقدير الذي يحصل عليه الطالب في الجامعة وجنسه أخذت عينة من نتائج الطلاب الذكور و الإناث وكانت كما يلي:

أولا: الإناث

راسب	راسب	راسب	جيد جدا	ممتاز	مقبول	ممتاز	جيد	جيد جدا	
جيد جدا	جيد	مقبول	مقبول	مقبول	راسب	مقبول	راسب	مقبول	
مقبول	راسب	جيد جدا	جيد جدا	جيد	جيد	ممتاز	جيد جدا	ممتاز	جيد
جيد	جيد	جيد جدا	ممتاز	جيد					

ثانيًا :الذكور

راسب	جيد	جيد	جيد	راسب	جيد جدا	راسب	جيد جدا	جيد جدا	جيد
راسب	راسب	راسب	راسب	راسب	مقبول	مقبول	ممتاز	ممتاز	راسب
راسب	مقبول	مقبول	ممتاز	ممتاز	جيد	جيد	مقبول	راسب	راسب
جيد	جيد	جيد جدا	ممتاز	جيد	جيد جدا	ممتاز			

والمطلوب هل توجد علاقة بين تقدير الطالب وجنسه عند مستوى الدلالة 0.05؟

اختبار مان-ويتني Mann Whitney Test

يستخدم هذا الاختبار عندما لا تتوافر شروط استخدام اختبار t لعينتين مستقلتين، أو أن تكون البيانات المتوفرة هي رتب القيم وليست قيمها، مما نضطر لاستخدام اختبار مان ويتني، حيث أنه يعتمد على رتب القيم.

تطبيق: لإجراء اختبار لمجموعة كبيرة من الطلبة، قام المدرس بوضع مجموعتين من الأسئلة أعطى المجموعة الأولى للطلبة الذين يجلسون على المقاعد ذات الأرقام الفردية، والمجموعة الثانية للطلبة الذين يجلسون على

المقاعد ذات الأرقام الزوجية. هل مجموعتا الأسئلة متكافئة أم لا؟ ولاختبار هذه الفرضية، قام المدرس برصد بعض العلامات من كل من المجموعتين وكانت العلامات كما يلي:

المجموعة (1)	52	78	56	90	65	86	64	90	49	78
المجموعة (2)	72	62	91	88	90	74	98	80	81	71

اختبار كروسكال– واليس ***Kruskal– Wallis***

يعتبر اختبار كروسكال واليس تعميم لاختبار مان ويتني وهو يستخدم لمقارنة توزيع 3 عينات فأكثر من المجتمعات المستقلة، وعادة ما يطبق عندما لا تتوافر شروط تطبيق اختبار تحليل التباين الأحادي أو عندما تكون البيانات المتوافرة هي بيانات ترتيبية وهو يعمل على:

- اختبار الفروق بين وسيط 3 عينات فأكثر.
- اختبار الفروق في طرق المعالجة.
- اختبار هل المجتمعات قيد الدراسة لها نفس التوزيع.
- اختبار الفروق في متوسطات المجتمعات (إن تعذر استخدام الطرق المعلمية).

تطبيق: في دراسة لمقارنة هل تعتمد درجات الطلاب على التخصص في مقرر مبادئ الإحصاء قام مدرس بإعطاء امتحان لمجموعة من الطلاب في المقرر ومن ثم رصد مجموعة من الدرجات وكانت الدرجات حسب التخصص كما يلي:

اقتصاد						تجارة							تسيير						
45	30	35	65	55	75	70	80	75	40	70	65	70	95	40	40	31	50	90	80
65	60	62	65	80	90	74	78	65	55	45	35	55	90	75	98	75	65	42	85
					68							30		95	90	77	40	85	85

والمطلوب هل توجد فروق بين توزيع درجات الطلاب في التخصصات الثلاثة أعلاه عند مستوى الدلالة 5%؟

اختبار ولكوكسن Wilcoxon test

اختبار ولكوكسن هو اختبار لامعلمي، يستخدم لدراسة الفروق بين متوسطي أو توزيعي عينتين مرتبطتين، بمعنى آخر هل يوجد اختلاف في توزيع (متوسط) بيانات العينتين أم لا.

تطبيق: يعطي الجدول التالي عدد الأميال التي تقطعها كل من 12 سيارة باستخدام نوعين من الوقود A و B.

A	**26.4**	**10.3**	**15.8**	**16.5**	**32.5**	**8.3**	**22.1**	**30.1**	**12.9**	**12.6**	**27.3**	**9.4**
B	**24.3**	**9.8**	**16.9**	**17.2**	**30.5**	**7.9**	**22.4**	**28.6**	**13.1**	**11.6**	**25.5**	**8.6**

والمطلوب: هل يختلف متوسط الأميال التي تقطعها السيارة باستخدام الوقود A عن متوسط الأميال التي تقطعها باستخدام الوقود B عند مستوى الدلالة 5 %؟

اختبار فريدمان Friedman Test

هو اختبار لامعلمي لثلاث عينات مرتبطة أو أكثر، وهو بديل لاختبار تحليل التباين المعلمي وهو من النوع: Repeated Measure Design ويطبق في حالة عدم توفر شروط تطبيق الاختبارات المعلمية وهو يعمل على المقارنة من حيث:

- الفروق في طرق المعالجة
- هل المجتمعات قيد الدراسة لها نفس التوزيع.
- اختبار الفروق في متوسطات المجتمعات.

تطبيق: لنفرض البيانات التالية

الرقم	1	2	3	4	5	6	7	8	9	10	11	12
T1	10	12	15	13	15	12	11	13	15	7	12	10
T2	18	19	17	14	20	15	7	18	19	13	13	8
T3	7	8	16	12	17	10	6	11	11	12	18	5

المطلوب: هل يوجد فرق في طرق المعالجة المبينة في الجدول أعلاه عند مستوى الدلالة 0.05 ؟

3- تصحيح تطبيقات الاختبارات المعلمية واللامعلمية

اختبار التوزيع الطبيعي – كولموغوروف سميرنوف Kolmogorov-Smirnov Test

تطبيق:

تمثل البيانات التالية درجات 30 طالبا في مقرر "الأدوات الإحصائية لتحليل البيانات":

21	32	76	82	90
40	30	65	92	80
88	45	82	60	70
89	89	80	70	90
92	88	90	50	60
85	77	92	65	76
79	86	86	79	68
31	90	71	82	94
29	94	93	68	83
50	97	68	80	74

المطلدوب: استخدم اختبار كولموغوروف – سميرنوف لمعرفة أن البيانات السابقة لها توزيع طبيعي أم لا مستخدما مستوى الدلالة $\alpha = 0.05$.

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية H_0: درجات الطلبة في مقرر الأدوات الإحصائية لتحليل البيانات تتوزع توزيعا طبيعيا عند المستوى5%.

الفرضية H_1: درجات الطلبة في مقرر الأدوات الإحصائية لتحليل البيانات لا تتوزع توزيعا طبيعيا عند المستوى 5%

.

ثانيا: مخرجات الاختبار من برنامج SPSS

نتبع المسار التالي: **Analyze ⇒ Descriptive Statistics ⇒ Explore**

Tests of Normality

	Kolmogorov-Smirnov[a]			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Y	,160	50	,003	,866	50	,000

a. Lilliefors Significance Correction

النتيجة الموضحة في الجدول الأول تبين أن قيمة إحصاءة اختبار كولموغوروف – سميرنوف $(K.S = 0.160)$ وأن القيمة الاحتمالية هي$sig = 0{,}003$ وهي أصغر من مستوى الدلالة $\alpha = 0{,}05$ لذلك نرفض الفرضية الصفرية ونقبل الفرضية البديلة لها القائلة بأن درجات طلاب مقرر (مساق) الأدوات الإحصائية لتحليل البيانات، لا تتبع التوزيع الطبيعي.

اختبار t للعينة الواحدة One Sample T- Test

تطبيق: البيانات التالية تمثل درجات عشرين طالبا في مساق مادة الرياضيات:

65 72 68 82 45 92 87 85 90 60 48 60 68 72 79 68 73 69 78 84

المطلوب: هل يختلف تحصيل هؤلاء الطلبة في مادة الرياضيات عن الدرجة 62 عند مستوى الدلالة 5%؟

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية H_0: لا يختلف تحصيل الطلبة في مادة الرياضيات عن الدرجة 62 عند مستوى الدلالة 5%؟

الفرضية H_1: يختلف تحصيل الطلبة في مادة الرياضيات عن الدرجة 62 عند مستوى الدلالة 5%؟

ثانيا: مخرجاتSPSS لاختبار t لعينة واحدة

نتبع المسار التالي: **Analyze ⇒ Compare Means ⇒ One-Sample t Test**

One-Sample Statistics

	N	Mean	Std. Deviation	Std. Error Mean
D-Math	20	72,2500	12,86724	2,87720

بقراءة بسيطة للجدول الأول يتضح أن حجم عينة الطلبة هو 20 وأن المتوسط الحسابي لتحصيلهم في مادة الرياضيات قد بلغ 72.25 بانحراف معياري 12.86724 وأن الخطأ المعياري للمتوسط هو 2.87720.

One-Sample Test للعينة الواحدة t اختبار

	Test Value = 62					
					95% Confidence Interval of the Difference	
	t	df	Sig. (2-tailed)	Mean Difference	Lower	Upper
D-Math	3,562	19	,002	10,25000	4,2279	16,2721

من خلال الجدول الثاني يتبين أن قيمة إحصاءة اختبار t للعينة الواحدة قد بلغت $(t = 3.562)$ وأن القيمة الاحتمالية لها $(sig = p - value = 0.002)$ وهي أصغر من مستوى الدلالة 5%، وهذا يعني أننا نرفض الفرضية الصفرية ونقبل الفرضية البديلة لها التي تنص على أن تحصيل الطلبة في الرياضيات يختلف جوهريا، وبشكل إيجابي، عن المحك (الدرجة 62).

وبطريقة أخرى: اعتمادا على مجال الثقة بدرجة 95% نلاحظ أن أصغر قيمة هي $(Lower = 4.2779)$ وأن أكبر قيمة $(UPPER = 16.2721)$ وهما من نفس الإشارة، إذن الصفر لا ينتمي إلى مجال الثقة $[4.2779; 16.2721]$، مما يعني نرفض الفرضية الصفرية ونقبل الفرضية البديلة لها، أي أن هناك اختلاف جوهري بين تحصيل الطلبة ودرجة المحك 62.

اختبار (t) لعينتين مستقلتين Independent Samples t- test

تطبيق: رغب أستاذ في تطوير مستوى الطالب من خلال معرفة أي الطريقتين أفضل، الطريقة التقليدية التي تعتمد على 30% في المخبر، أم الطريقة الجديدة التي تعتمد على 70% في المخبر. ولغرض التوصل إلى قرار مناسب اختار عينتين A و B بشكل مستقل وبعد مدة أجرى امتحان موحد لهما فكانت النتائج على النحو التالي:

A	85	70	60	75	65	75	78	77	65	80
B	70	85	90	65	75	85	80	88	82	

هل هناك فروق بين المجموعتين A و B في نتائج الامتحان عند مستوى دلالة 0.05؟

الحل: أولا: صياغة الفرضيات الإحصائية

الفرضية الصفرية: لا توجد فروق بين المجموعتين A و B في نتائج الامتحان عند مستوى دلالة 0.05

الفرضية البديلة: توجد فروق بين المجموعتين A و B في نتائج الامتحان عند مستوى دلالة 0.05

ثانيا: مخرجات SPSS لاختبار t لعينتين مستقلتين

نتبع المسار التالي: **Analyze ⇒ Compare Means ⇒ Independent-Samples- t- Test**

Group Statistics

	G	N	Mean	Std. Deviation	Std. Error Mean
Y	A	10	73,0000	7,80313	2,46757
	B	9	80,0000	8,42615	2,80872

الجدول الأول يوضح حجم العينتين $(N_A = 10.\ N_B = 9)$ والوسط الحسابي لكل منهما حيث $(\bar{X} = 73;\ \bar{X}_B = 80)$ والانحراف المعياري $(\sigma_A = 7.80313;\ \sigma_B = 842615)$ وأن الخطأ المعياري للمتوسط لكل من المجموعتين $(SE_A = 2.46757;\ SE_B = 2.80872)$

Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
								95% Confidence Interval of the Difference	
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	Lower	Upper
Y Equal variances assumed	,019	,893	-1,880	17	,077	-7,00000	3,72274	-14,85430	,85430
Equal variances not assumed			-1,872	16,420	,079	-7,00000	3,73869	-14,90921	,90921

الجدول الثاني يتضمن اختبارين، الأول هو اختبار ليفين (Levene) حيث قيمة الإحصاءة F=0.019 وأن القيمة الاحتمالية للاختبار هي $(sig = 0.893)$ وهي أكبر من مستوى الدلالة $\alpha = 0.05$، مما يعني أننا نقبل الفرضية الصفرية التي تنص على أن تباييني العينتين متساويان، وهذا يتطلب التعامل مع السطر الأول بالنسبة لاختبار t حيث $(t = -1.880)$ وأن القيمة الاحتمالية هي $(sig = 0.077)$ وهي أكبر من مستوى الدلالة $\alpha = 0.05$، مما يعني قبول الفرضية الصفرية التي تنص على أنه لا توجد فروق بين المجموعتين A و B في نتائج الامتحان عند مستوى دلالة 5%.

اختبار (t) لعينتين مرتبطتين Paired Samples t- test

تطبيق: تم إشراك 14 موظفا في دورة تكوينية وقيست معلوماتهم قبل وبعد اشتراكهم في الدورة، وكانت النتائج كالتالي:

رقم الموظف	الدرجة قبل الدورة	الدرجة بعد الدورة	رقم الموظف	الدرجة قبل الدورة	الدرجة بعد الدورة
1	209	217	8	217	230
2	215	219	9	231	240
3	215	210	10	239	242
4	227	232	11	223	214
5	22	226	12	228	227
6	212	219	13	223	235
7	234	238	14	223	226

المطلوب: هل للدورة التكوينية تأثير على مستوياتهم المعلوماتية عند مستوى دلالة 5%؟

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية لصفرية: لا ليس للدورة التكوينية تأثير على مستويات الموظفين المعلوماتية عند مستوى دلالة 5%؟

الفرضية البديلة: نعم للدورة التكوينية تأثير على مستويات الموظفين المعلوماتية عند مستوى دلالة 5%؟

ثانيا: مخرجات SPSS لاختبار t لعينتين مرتبطتين

نتبع المسار التالي: **Analyze ⇒ Compare Means ⇒ Paired-Samples- t- Test**

Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	before	2,0129E2	14	57,21945	15,29254
	after	2,2679E2	14	10,00137	2,67298

الجدول الأول يوضح مؤشرات العينة قبل الدورة التكوينية حيث كانت على النحو التالي:

$$(N_b = 14;\ \bar{X}_b = 201.29;\ \ \sigma_b = 57.21945;\ SE_b = 2.80872)$$

في حين أن النتائج أصبحت بعد الدورة التكوينية على النحو التالي:

$$(N_a = 14;\ \bar{X}_b = 226.79;\ \sigma_a = 10.00137;\ SE_a = 2.67298)$$

Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	before & after	14	-,009	,975

أما الجدول الثاني فيتعلق باختبار معامل ارتباط حيث $r = 0.009$ وأن القيمة الاحتمالية $(sig = 0.975)$ أكبر من مستوى الدلالة 0.05، مما يعني أننا نقبل الفرضية الصفرية التي تنص على أنه لا يوجد فرق جوهري بين العينتين.

Paired Samples Test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference Lower	Upper			
Pair 1 before - after	-2,55000E1	58,17844	15,54884	-59,09123	8,09123	-1,640	13	,125

أما النتيجة المهمة هنا فهي نتيجة اختبار t لعينتين مرتبطتين التي تظهر في الجدول الثالث، حيث نلاحظ أن قيمة إحصاءة الاختبار في $(t = -1.640)$ وأن القيمة الاحتمالية $(sig = 0.125)$ وهي أكبر من مستوى الدلالة 0.05، مما يعني أننا نقبل الفرضية الصفرية التي تنص على أن الدورة التكوينية لم تحدث تغييرا في مستويات الطلبة المعلوماتية عند مستوى المعنوية . 5%

تحليل التباين الأحادي باتجاه واحد One-Way ANOVA

تطبيق: استخدمت أربع طرق صناعية لإنتاج نوع معين من القماش بثلاث مكررات لكل طريقة وكانت النتائج حسب الجدول التالي:

المكررات / الطريقة	1	2	3	المتوسط
الطريقة 1	55	57	48	50
الطريقة 2	55	64	64	61
الطريقة 3	55	49	52	52
الطريقة 4	50	44	41	45

المطلوب: – هل توجد فروق جوهرية بين متوسطات الطرق الصناعية لإنتاج القماش عند مستوى دلالة 5%؟ في حالة ظهور فروق معنوية بين الطرق الصناعية، اختبر معنوية الفروق بين متوسطي كل طريقتين باستخدام طريقة الفرق المعنوي الأصغر $(L.S.D)$ لمستوى الدلالة 5%.

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية الصفرية: تساوي متوسطات الطرق الصناعية، أي $H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4$

الفرضية البديلة: عدم تساوي متوسطين على الأقل، أي $H_1: \exists\, \mu_i, \mu_j: \mu_i \neq \mu_j$

ثانيا: مخرجات SPSS لاختبار تحليل التباين

نتبع المسار التالي: **Analyze ⇒ Compare Means ⇒ One-Way ANOVA**

ANOVA

product

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	402,000	3	134,000	7,053	,012
Within Groups	152,000	8	19,000		
Total	554,000	11			

1) يبين جدول ANOVA أن قيمة إحصاءة الاختبار $F = 7.053$ وأن القيمة الاحتمالية لهذا الاختبار $(sig = 0{,}012)$ وهي أصغر من مستوى الدلالة 5%، مما يعني أننا ترفض الفرضية الصفرية ونقبل الفرضية البديلة لها والتي تنص على أنه يوجد على الأقل متوسطين غير متساويين.

2) نظرا لوجود فروق معنوية بين الطرق الصناعية، نقوم بالمقارنات المتعددة (الاختبارات البعدية) باستخدام طريقة $(L.S.D)$ بإتباع المسار التالي:

Analyze ⇒ Compare Means ⇒ One-Way ANOVA ⇒ *Post Hoc*

Multiple Comparisons المقارنات المتعددة

product
LSD

(I) method	(J) method	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
1	2	-11,000*	3,559	,015	-19,21	-2,79
	3	-2,000	3,559	,590	-10,21	6,21
	4	5,000	3,559	,198	-3,21	13,21
2	1	11,000*	3,559	,015	2,79	19,21
	3	9,000*	3,559	,035	,79	17,21
	4	16,000*	3,559	,002	7,79	24,21
3	1	2,000	3,559	,590	-6,21	10,21
	2	-9,000*	3,559	,035	-17,21	-,79
	4	7,000	3,559	,085	-1,21	15,21
4	1	-5,000	3,559	,198	-13,21	3,21
	2	-16,000*	3,559	,002	-24,21	-7,79
	3	-7,000	3,559	,085	-15,21	1,21

*. The mean difference is significant at the 0.05 level.

يبين جدول (...) المقارنات المتعددة أنه باستعمال طريقة $(S.D.)$ ظهرت فروق معنوية عند مستوى الدلالة 5% بين متوسطات المعالجات $(1 و 2)$، $(2 و 3)$، $(2 و 4)$ لأن القيمة الاحتمالية في هذه الحالات الثلاث جاءت أصغر من مستوى الدلالة 5%

ملاحظة: لاختبار التجانس (تساوي التباينات) نتبع المسار التالي:

Analyze ⇒ Compare Means ⇒ One-Way ANOVA ⇒ *Options*

Test of Homogeneity of Variances

product

Levene Statistic	df1	df2	Sig.
,667	3	8	,596

من خلال الجدول الأخير نلاحظ أن القيمة الاحتمالية $(sig = 0.596)$ وهي أكبر من مستوى الدلالة 5%، إذن نقبل الفرضية الصفرية التي تنص على تجانس التباين.

الاختبارات اللامعلمية (Nonparametric Tests)

اختبار جودة المطابقة Goodness of fit test لعينة واحدة

تطبيق: تمثل البيانات التالية عدد الأشخاص الذين تناولوا طعام العشاء في مطعم صغير على مدى 50 يومًا:

25	7	10	8	16	24	22	8	12	10	5	14	27	
15	20	12	16	19	24	6	10	1	15	23	8	30	
19	16	8	6	9	7	12	14	19	22	20	16	14	
		20	21	16	18	12	16	23	20	4	17	27	

المطلوب: هل متغير عدد الأشخاص الذين تناولوا العشاء في المطعم يتبع التوزيع الطبيعي عند مستوى الدلالة 5%؟

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية الصفرية: متغير عدد الأشخاص الذين تناولوا العشاء في المطعم يتبع التوزيع الطبيعي

الفرضية البديلة: عدد الأشخاص الذين تناولوا العشاء في المطعم لا يتبع التوزيع الطبيعي.

ثانيا: مخرجات SPSS لاختبار جودة المطابقة

نتبع المسار التالي: **Analyze ⇒ Nonparametric test ⇒ One-Sample K.S**

One-Sample Kolmogorov-Smirnov Test

		Dinner
N		50
Normal Parameters[a,b]	Mean	15.26
	Std. Deviation	6.782
Most Extreme Differences	Absolute	.081
	Positive	.081
	Negative	-.069
Kolmogorov-Smirnov Z		.573
Asymp. Sig. (2-tailed)		.898

a. Test distribution is Normal.

حجم العينة

متوسط البيانات

الانحراف المعياري للبيانات

أكبر فرق بين البيانات ودالة التوزيع الاحتمالية

قيمة اختبار جودة المطابقة

مستوى دلالة الاختبار

تبين النتائج أعلاه أن متوسط عدد الزبائن هو 15.26 بانحراف معياري قدره 6.782 وأن قيمة اختبار جودة المطابقة هو 0.573 وأن القيمة الاحتمالية *Sig= 0.898* وهي أكبر من مستوى الدلالة 5%. إذن نقبل الفرضية الصفرية التي تنص على أن البيانات تتبع التوزيع الطبيعي.

اختبار كاي تربيع Chi-Square test للاستقلالية

تطبيق: في دراسة للعلاقة بين التقدير الذي يحصل عليه الطالب في الجامعة وجنسه أخذت عينة من نتائج الطلاب الذكور و الإناث وكانت كما يلي:

أولا: الإناث

راسب	راسب	راسب	راسب	جيد جدا	ممتاز	مقبول	ممتاز	جيد	جيد جدا
جيد جدا	جيد	مقبول	مقبول	مقبول	راسب	مقبول	راسب	مقبول	مقبول
مقبول	راسب	جيد جدا	جيد جدا	جيد	جيد	ممتاز	جيد جدا	ممتاز	جيد
جيد	جيد	جيد جدا	ممتاز	جيد					

ثانيًا :الذكور

راسب	جيد	جيد	جيد	راسب	جيد جدا	راسب	جيد جدا	جيد جدا	جيد
راسب	راسب	راسب	راسب	راسب	مقبول	مقبول	ممتاز	ممتاز	راسب
راسب	مقبول	مقبول	ممتاز	ممتاز	جيد	جيد	مقبول	راسب	راسب
جيد	جيد	جيد جدا	ممتاز	جيد	جيد جدا	ممتاز			

والمطلوب: هل توحد علاقة بين تقدير الطالب وجنسه عند مستوى الدلالة 0.05؟

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية الصفرية: لا يعتمد تقدير الطالب على جنسه(متغير الجنس والتقدير مستقلان)

الفرضية البديلة: يعتمد تقدير الطالب على جنسه (توجد علاقة بين جنس الطالب وتقديره)

ثانيا: مخرجات SPSS لاختبار كاي تربيع للاستقلالية

نتبع المسار التالي: **Analyze ⇒ Descriptive Statistics ⇒ Cross tabs**

تتكون نتائج الأمر *Cross tabulation* من ثلاثة جداول، حسب ما يلي:

Case Processing Summary

	Cases					
	Valid		Missing		Total	
	N	Percent	N	Percent	N	Percent
Result * Gender	72	100,0%	0	,0%	72	100,0%

الجدول الأول يصف حجم العينات المدخلة ونسب البيانات المفقودة.

أما الجدول الثاني فيبين أن عدد البيانات المدخلة 72 ، عدد الذكور 37 (منهم 12 راسب وقيمتها المتوقعة 9.76، 5 مقبول وقيمتها المتوقعة 6.68 ، 9 جيد وقيمتها المتوقعة 8.74 ، 5 جيد جدا وقيمتها المتوقعة 6.17 ، و 6 ممتاز وقيمتها المتوقعة 5.65) والإناث 35 (منهم 7 راسب وقيمتها المتوقعة 9.24 ، 8 مقبول وقيمتها المتوقعة 6.32 ، 8 جيد وقيمتها المتوقعة 8.26 ، 7 جيد جدا وقيمتها المتوقعة 5.83 ، و 5 ممتاز وقيمتها المتوقعة 5.35).

Result * Gender Crosstabulation

			Gender		
			male	female	Total
Result	failure	Count	12	7	19
		Expected Count	9,8	9,2	19,0
	accepted	Count	5	8	13
		Expected Count	6,7	6,3	13,0
	good	Count	6	7	13
		Expected Count	6,7	6,3	13,0
	very god	Count	8	8	16
		Expected Count	8,2	7,8	16,0
	excellent	Count	6	5	11
		Expected Count	5,7	5,3	11,0
	Total	Count	37	35	72
		Expected Count	37,0	35,0	72,0

ويبين الجدول الثالث أن قيمة اختبار مربع كاي هي 2.437 بدرجة حرية $df = 4$ ، في حين أن القيمة الاحتمالية للاختبار $(sig = 0.656)$ وهي أكبر من مستوى الدلالة 0.005، إذن نقبل الفرضية الصفرية التي تنص على أن تقدير الطالب لا يعتمد على جنسه.

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)
Pearson Chi-Square	2,122[a]	4	,713
Likelihood Ratio	2,142	4	,710
Linear-by-Linear Association	,145	1	,704
N of Valid Cases	72		

a. 0 cells (,0%) have expected count less than 5. The minimum expected count is 5,35.

اختبار مان–ويتني Mann Whitney Test

تطبيق:

الموضوع الرابع: الاختبارات المعلمية والاختبارات اللامعلمية

لإجراء اختبار لمجموعة كبيرة من الطلبة، قام المدرس بوضع مجموعتين من الأسئلة أعطى المجموعة الأولى للطلبة الذين يجلسون على المقاعد ذات الأرقام الفردية، والمجموعة الثانية للطلبة الذين يجلسون على المقاعد ذات الأرقام الزوجية. هل مجموعتا الأسئلة متكافئة أم لا؟ ولاختبار هذه الفرضية، قام المدرس برصد بعض العلامات من كل من المجموعتين وكانت العلامات كما يلي:

المجموعة (1)	52	78	56	90	65	86	64	90	49	78
المجموعة (2)	72	62	91	88	90	74	98	80	81	71

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية H_0: مجموعتا الأسئلة متكافئة (لا توجد فروق ذات دلالة إحصائية بين مجموعتي الأسئلة) عند5%.

الفرضية H_1: مجموعتا الأسئلة غير متكافئة (توجد فروق ذات دلالة إحصائية بين مجموعتي الأسئلة) عند5%.

ثانيا: مخرجات SPSS لاختبار مان ويتني

نتبع المسار التالي: **Analyze ⇒ Nonparametric test ⇒ 2 Independent Samples**

Ranks الرتب

	group	N	Mean Rank	Sum of Ranks
degree	1	10	8,60	86,00
	2	10	12,40	124,00
	Total	20		

يصف الجدول الأول العينتين حيث يظهر لنا أن حجم عينة المجموعة الأولى 10 ومتوسط رتبهم 8.6 ومجموع هذه الرتب 86 ، أما بالنسبة للمجموعة الثانية فحجم العينة 10 ومتوسط الرتب 12.4 بمجموع قدره 124وأن حجم العينة الكلية 20.

Test Statistics[b]

	degree
Mann-Whitney U	31,000
Wilcoxon W	86,000
Z	-1,439
Asymp. Sig. (2-tailed)	,150
Exact Sig. [2*(1-tailed Sig.)]	,165[a]

a. Not corrected for ties.

b. Grouping Variable: group

يبين الجدول الثاني أن قيمة اختبار مان ويتني هي 31 وأن القيمة الاحتمالية $(sig = 0{,}150)$ وهي أكبر من مستوى الدلالة 5%. إذن نقبل الفرضية الصفرية التي تنص على أن مجموعتي الأسئلة متكافئة، أي لا توجد فروق ذات دلالة إحصائية بين مجموعتي الأسئلة عند الدلالة مستوى5%.

اختبار كروسكال– واليس ***Kruskal– Wallis***

تطبيق: في دراسة لمقارنة هل تعتمد درجات الطلاب على التخصص في مقرر مبادئ الإحصاء قام مدرس بإعطاء امتحان لمجموعة من الطلاب في المقرر ومن ثم رصد مجموعة من الدرجات وكانت الدرجات حسب التخصص كما يلي:

اقتصاد	تجارة	تسيير

95 40 40 31 50 90 80	70 80 75 40 70 65 70	45 30 35 65 55 75
90 75 98 75 65 42 85	74 78 65 55 45 35 55	65 60 62 65 80 90
95 90 77 40 85 85	30	68

والمطلوب: هل توجد فروق بين توزيع درجات الطلاب في التخصصات الثلاثة أعلاه عند مستوى الدلالة 5%؟

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية H_0: لا توجد فروق في وسيط توزيع درجات الطلاب في الإحصاء في التخصصات الثلاثة **عند** 5%

الفرضية H_1: توجد فروق في وسيط توزيع درجات الطلاب في الإحصاء في التخصصات الثلاثة 5%

ثانيا: مخرجات SPSS لاختبار مان ويتني

نتبع المسار التالي: **Analyze ⇒ Nonparametric test ⇒ K Independent Samples**

تتكون مخرجات الاختبار من جدولين: الأول خاص بوصف نتائج العينات، والثاني خاص باختبار كروسكال – واليس نفسه

Ranks الرتب

	specjality	N	Mean Rank
Degree	economy	13	20,69
	commerce	15	20,77
	gestion	20	29,78
	Total	48	

يبين الجدول الأول المعنون بـ: *Ranks*، حجم العينات $(n_1 = 13,\ n_2 = 15,\ n_3 = 20)$ ومتوسط رتب كل عينة حيث يبلغ متوسط رتب تخصص الاقتصاد $\overline{X}_1 = 20.69$ أما متوسط رتب تخصص التجارة فهو $\bar{X}_2 = 20.77$ ، في حين أن متوسط رتب تخصص التسيير قد بلغ $\bar{X}_3 = 29.78$. أما مجموع حجوم العينات يساوي50.

Test Statistics[a,b]

	Degree
Chi-Square	4,890
df	2
Asymp. Sig.	,087

a. Kruskal Wallis Test

b. Grouping Variable: specialty

ويبين جدول النتائج الثاني أن قيمة اختبار كروسكال–واليس $(K.W = \chi^2 = 4.374)$ بدرجة حرية $df = 2$وأن القيمة الاحتمالية $sig = 0.087$ وهي أكبر من مستوى الدلالة 5%. إذن نقبل الفرضية الصفرية التي تنص على أنه لا توجد فروق ذات دلالة إحصائية في وسيط توزيع درجات الطلاب في مقرر الإحصاء ضمن التخصصات الثلاثة عند مستوى الدلالة 5%.

ملاحظة هامة: في حالة وجود فروق معنوية فقط، نقوم بالاختبارات البعدية من أجل تحديد اتجاهاتها كما هو مبين في اختبار تحليل التباين الأحادي باتجاه واحد.

اختبار ولكوكسن Wilcoxon test

تطبيق: يعطي الجدول التالي عدد الأميال التي تقطعها كل من 12 سيارة باستخدام نوعين من الوقود A و B.

A	**26.4**	**10.3**	**15.8**	**16.5**	**32.5**	**8.3**	**22.1**	**30.1**	**12.9**	**12.6**	**27.3**	**9.4**
B	**24.3**	**9.8**	**16.9**	**17.2**	**30.5**	**7.9**	**22.4**	**28.6**	**13.1**	**11.6**	**25.5**	**8.6**

والمطلوب: هل يختلف متوسط الأميال التي تقطعها السيارة باستخدام الوقود A عن متوسط الأميال التي تقطعها باستخدام الوقود B عند مستوى الدلالة 5 %؟

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية الصفرية: لا يختلف متوسط الأميال التي تقطعها السيارة باستخدام الوقود A عن متوسط الأميال التي تقطعها باستخدام الوقود B عند مستوى الدلالة 5 %؟

الفرضية البديلة: يختلف متوسط الأميال التي تقطعها السيارة باستخدام الوقود A عن متوسط الأميال التي تقطعها باستخدام الوقود B عند مستوى الدلالة 5 %؟

ثانيا: مخرجات SPSS لاختبار ولكوكسن

نتبع المسار التالي: **Analyze ⇒ Nonparametric test ⇒ 2 Related Samples**

تتكون مخرجات SPSS لاختبار ولكوكسن من الجدولين التاليين:

Ranks الرتب

		N	Mean Rank	Sum of Ranks
B - A	Negative Ranks	8[a]	7,75	62,00
	Positive Ranks	4[b]	4,00	16,00
	Ties	0[c]		
	Total	12		

a. B < A

b. B > A

c. B = A

يبين الجدول الأول (بعنوانRanks) أن حجم العينة الكلي هو 12 منهم 8 حالات كانت فيها رتب قيم المتغير B أقل من رتب قيم A ويرمز لهل بالرمز $a.B < A$ أو وتسمى الرتب السالبة، وأن متوسط رتب هذه الحالات قد بلغ 7.75 وبمجموع 62 ، ويبين الجدول أيضا أن هناك 4 حالات كانت فيها رتب قيم المتغير B أكبر من رتب قيم A ويرمز لها بالرمز $B > A$ وتسمى الرتب الموجبة، وأن متوسط هذه الرتب قد بلغ 4 وبمجموع 16، وأنه لا توجد حالات كانت رتب قيم المتغيرين متساوية ويرمز لها بالرمز $A = B$.

Test Statistics[b]

	B - A
Z	-1,804[a]
Asymp. Sig. (2-tailed)	,071

a. Based on positive ranks.

b. Wilcoxon Signed Ranks Test

كما يبين الجدول الثاني أن قيمة إحصاءة اختبار هي $Z = -1.804$، ودرجة حرية وأن القيم الاحتمالية المصاحبة لها $(sig = 0{,}071)$ وهي أكبر من مستوى الدلالة 5%، مما يعني أننا نرفض الفرضية الصفرية

ونقبل الفرضية البديلة لها والتي تنص على أن متوسط الأميال التي تقطعها السيارة باستخدام الوقود A يختلف عن متوسط الأميال التي تقطعها باستخدام الوقود B عند مستوى الدلالة 5 %؟

اختبار فريدمان Friedman Test

تطبيق: لنفرض البيانات التالية

الرقم	T_1	T_2	T_3
1	10	18	7
2	12	19	8
3	15	17	16
4	13	14	12
5	15	20	17
6	12	15	10
7	11	7	6
8	13	18	11
9	15	19	11
10	7	13	12
11	12	13	18
12	10	8	5

المطلوب: اختبر هل يوجد فرق في طرق المعالجة المبينة في الجدول أعلاه عند مستوى الدلالة 0.05 ؟

الحل:

أولا: صياغة الفرضيات الإحصائية

الفرضية الصفرية: لا يوجد فرق في طرق المعالجة عند مستوى الدلالة 0.05

الفرضية البديلة: يوجد فرق في طرق المعالجة عند مستوى الدلالة 0.05

ثانيا: مخرجات SPSS لاختبار فريدمان

نتبع المسار التالي: Analyze ⇒ **Nonparametric test ⇒ K Related Samples**

تتكون مخرجات SPSS لاختبار فريدمان من جدولين:

Ranks

	Mean Rank
T1	1,83
T2	2,75
T3	1,42

يبين الجدول الأول أعلاه أن متوسط رتب العينة الأولى هو 1.83 ومتوسط رتب العينة الثانية 2.75 أن متوسط رتب العينة الثالثة هو 1.42.

Test Statistics[a]

N	12
Chi-Square	11,167
df	2
Asymp. Sig.	,004

a. Friedman Test

ويبين الجدول الثاني أن حجم في كل عينة هو 12 وأن درجة الحرية هي $df = k - 1 = 2$ ، وأن قيمة اختبار فريدمان هي 11.167 وأن القيمة الاحتمالية $sig = 0{,}004$ وهي أصغر من مستوى الدلالة 0.05 ،

مما يعني أننا نرفض الفرضية الصفرية ونقبل الفرضية البديلة لها التي تنص على أنه يوجد فرق معنوي في طرق المعالجة عند مستوى الدلالة 5%.

ولتحديد اتجاهات الفروق بين المعالجات الثلاث، نقوم بالاختبارات البعدية كما هو معروف.