محاور مقياس إحصاء استدلالي:(السنة الثانية لسانس تربية حركية وتدريب رياضي)-السداسي الثاني:

1- تذكير بالإحصاء الوصفي /مقاييس النزعة المركزية ومقاييس التشتت
2- مدخل عام للإحصاء الاستدلالي
3- المجتمع الإحصائي والعينة
4- اختبار الفروض واتخاذ القرارات الإحصائية
5- اختبار الفروض حول متوسطات الحسابية لعينة واحدة(ت ستودنت)
6- اختبار الفروض حول متوسطات الحسابية لعينتين مرتبطتين (ت ستودنت)
7- اختبار الفروض حول متوسطات الحسابية لعينتين مستقلتين (ت ستودنت)
8- اختبار Z للعينات الكبيرة(متساوية وغير متساوية العدد)
9- اختبارات الارتباط(بيرسون، سبيرمان، فاي، الانحدار الخطي)
10- اختبار كاف تربيع
11- تحليل التباين لعامل واحد (ANOVA)
12- تحليل التباين لعاملين.

محاضرة رقم 01: تذكير بالإحصاء الوصفي

محاضرة رقم02: مدخل عام للإحصاء الاستدلالي

محاضرة 01:

مصطلحات إحصائية:

الإحصاء الوصفي Descriptive statistics: يركز على جمع البيانات وتبويبها وتلخيصها وعرضها باستخدام الجداول والأشكال البيانية التي تصف خصائص البيانات الرئيسية.

-الإحصاء الاستدلالي Inferential statistics: يبحث في تعميم النتائج من العينة إلى المجتمع بواسطة الإجابة على مجموعة من الأسئلة البحثية أو اختبار جملة من الفروض الإحصائية المصاغة سابقا.

المجتمعpopulation : يمثل كل الأفراد أو أعضاء المجموعة أو المشاهدات قيد الدراسة ويرمز له بالرمز N.

- الوحدة الإحصائية unité statistique: عبارة عن عنصر من عناصر المجتمع المدروس يحمل صفة أو عدة صفات.فالوحدة في لغة الإحصاء تعني كائنا متحركا أو جامدا قابلا للقياس والعد.كالعامل في مجتمع العمال، أو الطالب في مجتمع الطلبة، ويرمز لها بالرمز**U**، **مثال ذلك**: العامل في مجتمع العمال، الطالب في مجتمع الطلبة.
- العينةéchantillon:هي تجمع يضم عددا من الوحدات أو الأفراد ممثلة للمجتمع الإحصائي (أي أن تعكس كل الصفات الموجودة في المجتمع الإحصائي)، كما يشترط أن يختار أفرادها بطريقة عشوائية (أي أن يكون لكل فرد من أفراد المجتمع الإحصائي نفس الفرصة لأن يكون عنصرا من عناصر العينة)، ويرمز لها بالرمز**n**.

المعلمةparameter: هي عبارة عن مقياس كمي لوصف خصائص المجتمع مثل الوسط الحسابي للمجتمع والانحراف المعياري للمجتمع.

المعاينةsampling: هي الطريقة التي يتم من خلالها اختيار عينة من المجتمع بهدف تقدير معالم ذلك المجتمع من أجل تعميم إحصاءات العينة.(خير، أبو زيد، 2018، ص ص 23-24)

مستويات القياس: يشير القياس إلى القيم الرقمية المستخدمة في تسجيل المشاهدات، ومستوى القياس يحدد العلاقة بين القيم المخصصة لقيم المتغير والى الأنماط الموجودة في تلك العلاقة، وأهمية القياس تكمن في المساعدة على تفسير بيانات المتغير، إضافة إلى أن تحديد طبيعة الاختبار الإحصائي يعتمد أساسا على مستوى القياس، وبصورة عامة تنقسم مستويات القياس إلى أربعة مستويات هي: الاسمي Nominal الرتبي Ordinal الفئوي(متساوى المسافة)Equal interval والنسبي Ratio

-الاسمي Nominal: أقل مستويات القياس من حيث الدقة حيث أن القيم الرقمية لأنها مجرد رموز مختصرة للإشارة إلى أسماء الفئات كما أن ترتيب تلك الفئات عشوائي وليس له معنى من حيث الأفضلية أو من حيث الأكثرية أو الأقلية

ومثال ذلك الفئات: مسلم، مسيحي، يهودي، يخصص لها القيم (1)، (2)، (3) بطريقة عشوائية لتكون القيم الرقمية مجرد رموز(لا تحمل معنى يفيد الكم) مختصرة تشير إلى الكلمات التي وردت في الفئات، إلا أن الرقم الأكبر لا يعني انه الأهم كما أن الرقم الأصغر لا يعني انه أقل أهمية كما أن القيمة (2) لا تعني أنها مضاعف العدد (1)

-الرتبي Ordinal: يتم في هذا المستوى ترتيب الفئات من الصغير إلى الكبير أو العكس، والمسافة بين الفئات تكون غير متساوية وليست ذات معنى، ومثال ذلك مستوى التعليم يتم ترتيبه إلى فئات من الأقل إلى الأعلى مثل ابتدائي، ثانوي، جامعي وفق القيم (1)، (2)، (3).

-متساوي المسافة(الفئوي)interval : ويسمى أيضا المستوى الفتري في هذا المستوى للمسافة بين الفئات أهمية ومعنى فمثلا عند قياس درجة الحرارة فان المسافة بين (15-25) هي نفس المسافة من (30-40) غير أن درجة الحرارة (40) لا تعبر عن ضعف درجة الحرارة عندما تكون (20).

-النسبةratio: هو اعلي مستويات القياس ويأخذ الصفر قيمة حقيقية التي تمثل بداية المقياس من الناحية النظرية للتعبير عن عدم وجود الشيء.(شراز،2015، ص ص 9-11)

في المستوى الاسمي لا يمكن إجراء العمليات الحسابية الأربع ويستخدم فيه المقاييس بالطرق اللامعلمية بغض النظر عن حجم العينة المنوال بدل الوسط الحسابي ، وفي المستوى الرتبي لا تستخدم أيضا العمليات الحسابية الأربع واغلب استخداماته في استمارات الاستبيان(ليكرت ثنائي، ثلاثي، رباعي...)ويتم استخدام الإحصاء اللامعلمي.

أما في المستوى الفتري تستخدم فيه عمليات الجمع والطرح وبالتالي إمكانية استخدام المتوسطات الحسابية والانحرافات المعيارية والإحصاء المعلمي هو الأنسب واهم ما يميز هذا المقياس الصفر فيه لا يعد مطلقا(غير حقيقي) مثل درجة الحرارة صفر لا يعني انعدام درجة الحرارة كما انه لا يمكن استخدام معامل الاختلاف، بينما مستوى النسبة يكون فيه الصفر دال على انعدام الحالة مثل وزن صفر بمعنى الوزن منعدم، ويمكن استخدام جميع العمليات الحسابية مع معامل الاختلاف إضافة إلى الإحصاء المعلمي إلا في حالات نادرة.(الحسني، وشامل جاسم، 2018)

-المتغير وأنواع المتغيرات:

المتغيرvariable: هو كل شيء يأخذ عدد من القيم المختلفة في الوقت نفسه ويمكن قياسه مثل الجنس، تحصيل الطلبة، أرباح الشركات، أو قد يكون متغيرا عبر الزمن مثل إدراك الفرد لظاهرة معينة عبر الزمن، أو مراحل مراض ما، ويمكن أن يحتل المتغير أي موضع من مقياس متصل مثل الطول والوزن، أو يخذ قيم محددة مثل عدد الأطفال في العائلة ويسمى في هذه الحالة متغيرا منفصلا.

أقسام المتغيرات: تقسم المتغيرات إلى قسمين رئيسيين:

1-المتغيرات الكيفية (النوعية)variables qualitatives: هذه المتغيرات وصفية و لا تأخذ فيها الأعداد معنى كمي مثل: اللغة، الجنس، الوظيفة. نلاحظ في هذا النوع من المتغيرات أن التصنيف يكون على أساس امتلاك الفرد للخاصية أو السمة أو عدم امتلاكه.

هذا النوع من المتغيرات يمكن أن يكون في مستوى قياس اسمي/ تصنيفي مثل: الديانة، اللون، الجنس، الحالة الاجتماعية (أعزب، متزوج، مطلق، أرمل). كما يمكن أن يكون له ترتيب أو تصنيف معين فيكون في مستوى ترتيبي مثل دخل الفرد (مرتفع، متوسط، منخفض) أو درجة مشاركة الطلبة في مقياس الإحصاء (كبيرة، متوسطة، ضعيفة)

أما فيما يتعلق بالمعالجة الإحصائية فيتم الاعتماد على الأساليب الإحصائية اللابارامترية

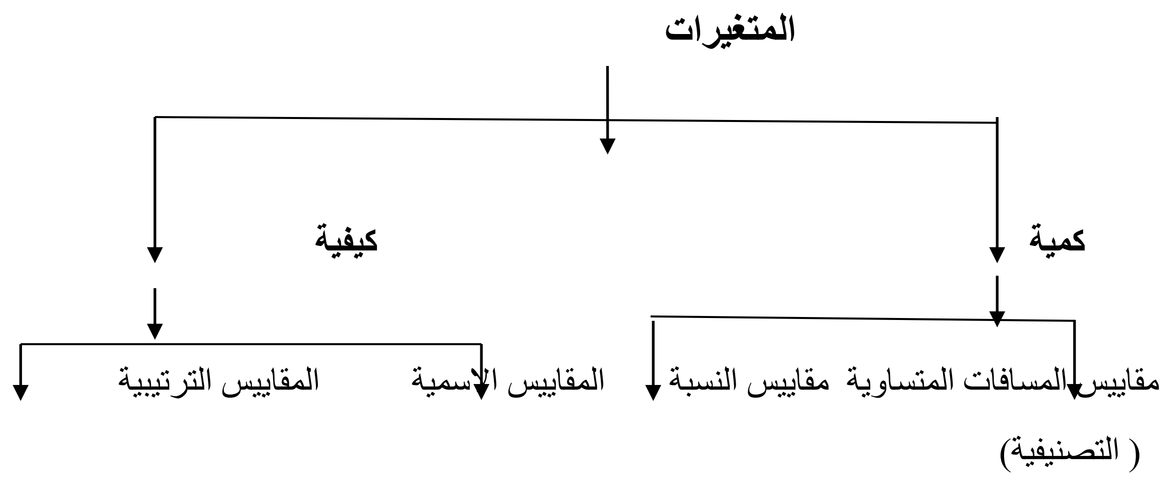
2-المتغيرات الكمية variables qualitatives: هي متغيرات نقاس بمقدار مثل: الوزن، الطول، السعة. ويقسم هذا النوع من المتغيرات إلى قسمين:

1-2- متغيرات كمية متصلة(مستمرة) variables continues: وهي تلك المتغيرات التي يمكن أن تأخذ قيمها أرقاما صحيحة أو كسرية مثل: درجة الحرارة، الوزن، الطول، العمر، الأجر. ونلاحظ هنا في هذا النوع من المتغيرات أنه يمكن تقسيم وحدات قياسه إلى وحدات جزئية بحيث تكون هناك استمرارية في القياس.

2-2- متغيرات كمية منفصلة(متقطعة) variables discrètes: وهي متغيرات نعبر عنها بأرقام عددية صحيحة مثل: عدد العمال، عدد المؤسسات الاقتصادية الخاصة، عدد الوفيات، عدد المساكن في حي من الأحياء...

إن هذه المتغيرات يمكن أن تندرج تحت مقياس المسافات المتساوية مثل: درجة الحرارة، درجة غليان الماء، عدد الوفيات، عدد العمال. كما يمكن أن تندرج تحت مقاييس النسبة مثل: الطول، الوزن...

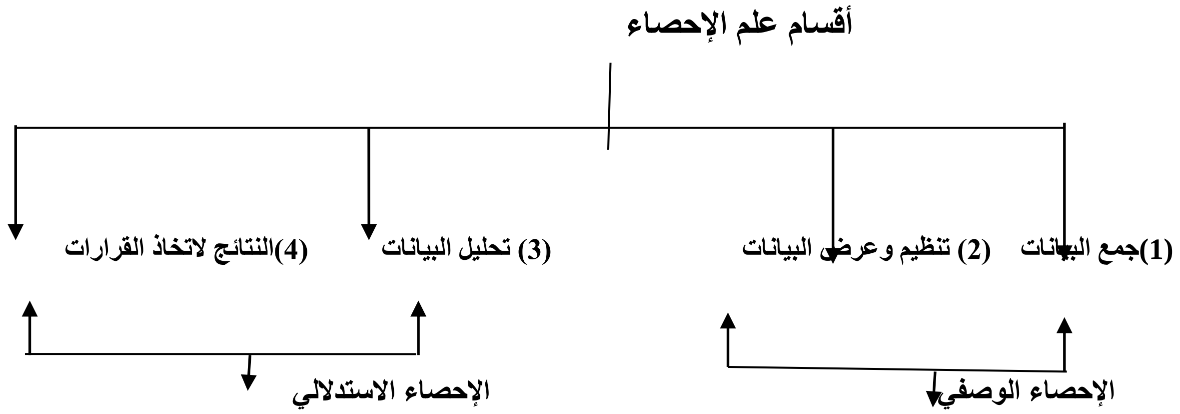
والشكل التالي يوضح أنواع المتغيرات والمقاييس المناسبة لهذه المتغيرات



أنواع الإحصاء	
الوصفي	**الاستدلالي**
-طرق تنظيم وتلخيص ووصف البيانات وصفا كميا للعينة المستخدمة. -مجموعة من المفاهيم والأساليب الإحصائية التي تستخدم في تنظيم وتلخيص وعرض مجموعة من البيانات بهدف إعطاء فكرة عامة عنها. -ملخص جيد لمجموعة كبيرة من المعلومات والبيانات. -أهم صورها التصنيف في جداول التوزيع التكراري والرسوم البيانية التي تعبر عن هذا التوزيع. -التلخيص له ثلاثة صور: *النزعة المركزية: المتوسط-الوسيط-المنوال *التشتت: المدى-الانحراف المعياري- نصف المدى الربيعي *العلاقة أو الارتباط والانحدار	مجموعة من الأساليب الإحصائية المستخدمة للتوصل إلى استنتاجات من بيانات العينة إلى المجتمع الأكبر -يشير إلى طرق الاستدلال عن المجتمع من بيانات العينة -عملية اتخاذ قرار منطقي باستخدام بيانات العينة وأسلوب إحصائي مناسب. -يعتمد على افتراضين أساسيين هما : العشوائية في اختيار العينة المستخدمة في الدراسة، والتوزيع الإعتدالي للمتوسطات. -ومنه اختبارات: تحليل التباين-اختبار مان ويتني-النسبة الحرجة-فريدمان-كروكسال واليس-ويلكوكسون-كا^2

المصدر: (محمود شعيب، ومحمود شعيب،2016، ص38)

الشكل الموالي يوضح أقسام الإحصاء:



محاضرة 02:

الإحصاء الاستدلالي يسمى كذلك الإحصاء الاستنتاجي أو الإحصاء التطبيقي وهو يعنى باختبار الفروض الإحصائية ويشتمل على نوعين من الإحصاء هما الإحصاء البرامتري والإحصاء اللابرامتري

والجدول التالي يوضح الفرق بين الأساليب البارامترية واللابارامترية

الأساليب البارامترية	الأساليب اللابارامترية

- تستخدم في التوزيعات الإعتدالية - تصلح للعينات الكبيرة غالبا - تناسب بيانات المسافات المتساوية والنسبة - يستغرق وقتا وجهدا - أكثر قوة	- تستخدم في التوزيعات الحرة - تصلح للعينات الصغيرة والكبيرة أحيانا - تناسب البيانات الاسمية وبيانات الرتبة وتصلح أحيانا للمسافات والنسبة - أسرع وأسهل استخداما - اقل قوة

المصدر: الإحصاء اللابارامتري (زكريا الشربيني2001ص100)

تصاغ الفروض الإحصائية في شكل صفري أو بديل.

1-الفرض الصفريH_0: يفترض الباحث أن العلاقة بين متغيرين أو الفرق بينهما يساوي صفر

مثال:

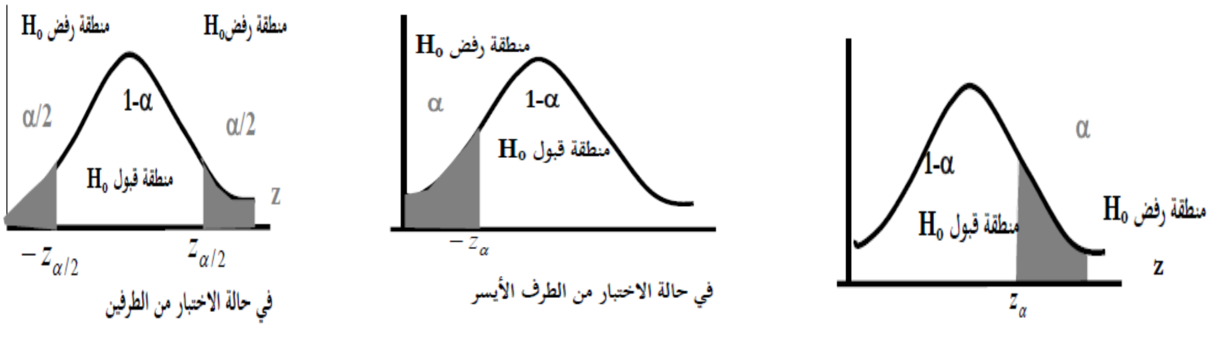
لا توجد علاقة إرتباطية ذات دلالة إحصائية بين متغيري الدافعية وطول الرياضي . $X_1-X_2=0$

- لا توجد فروق ذات دلالة إحصائية في الأداء الرياضي بين لاعبي كرة القدم ولاعبي كرة اليد $n_1-n_2=0$

2-الفرض البديلH_1:يفترض الباحث أن هناك علاقة بين متغيرين أو فروق متوقعة بينهم.

مثال:

- توجد علاقة إرتباطية ذات دلالة إحصائية بين متغيري الدافعية والأداء لدى لاعبي الكرة الطائرة.
- $X_1-X_2\neq0$
- توجد فروق ذات دلالة إحصائية في الثقة الرياضية بين الذكور والإناث..$n_1-n_2\neq0$
- وقد يكون الفرض البديل غير موجه ونستخدم فيه الإشارة($\neq$)بمعنى أننا لا نرجح كفة متوسط مجموعة على الاخرى في المقارنة وبيانيا يكون منحنى التوزيع الطبيعي ثنائي الطرف بمعنى منطقتين للرفض واحدة على اليمين والأخرى على اليسار، أما إذا كان الفرض البديل موجه فإننا نستخدم إحدى العلاقتين(< أو >) وفي هذه الحالة يكون شكل التوزيع الطبيعي يتضمن منطقة رفض واحدة إما على اليمين أو على اليسار.
- والأشكال الموالية توضح هذه الحالات بيانيا:



-مستوى الدلالة: هو المستوى الذي يطمئن عنده الباحث من صحة نتائجه وأنها لا تعود للصدفة. ويتم الكشف عنها من خلال جداول إحصائية خاصة وذلك بعد تحديد القيمة المحسوبة. وتكون هذه الجداول غالبا في ملاحق كتب الإحصاء.

وهناك ثلاث مستويات دلالة مقبولة إحصائيا:

1- مستوى دلالة 0.001، أي أن هناك ثقة في النتائج التي توصلت إليها كباحث بنسبة0.999 مقابل شك بنسبة 0.001 أي أن كل 1000مرة نقوم فيها بحساب معامل الارتباط مثلا، هناك 999 مرة صواب مقابل مرة واحدة محتملة للخطأ.
2- مستوى دلالة 0.01، أي أن هناك ثقة في النتائج التي توصلت إليها كباحث بنسبة0.99 مقابل شك بنسبة 0.01

أي أن كل 100مرة نقوم فيها بحساب معامل الارتباط ، هناك 99 مرة صواب مقابل مرة واحدة محتملة للخطأ.

3- مستوى دلالة 0.05، أي أن هناك ثقة في النتائج التي توصلت إليها كباحث بنسبة0.95 مقابل شك بنسبة0.05

أي أن كل 100مرة نقوم فيها بحساب معامل الارتباط ، هناك95مرة صواب مقابل 5 مرات محتملة للخطأ. ويعتبر هذا المستوى من الدلالة أقل مستوى نقبله كباحثين.

-اختبار الفروض الإحصائية: إن اختبار الفروض بأسلوب إحصائي يؤدي إلى اتخاذ قرار إذا ما كان الفرض مقبولا أم مرفوضا. تجدر الإشارة إلى أن قبول الفرض لا يعني بالضرورة أن يكون صحيحا، كما رفض الفرض لا يعني بالضرورة أن يكون خاطئا. والجدول التالي يوضح ذلك:

القرار / الفرضية	(H0) صحيح	(H0)خاطيء
قبول(H0)	**صواب $(1-\alpha)$**	**خطأ2(خطا من النوع الثاني β)**
رفض(H0)	**خطأ1(خطا من النوع الأول α**	**صواب $(1-\beta)$**

-اختبار فرض حول متوسط مجتمع:

للوصول إلى اتخاذ قرار بخصوص هذا الموضوع النسبة لاختبار فرض حول متوسط مجتمع μ أو نسبة في المجتمع P نتبع الخطوات التالية:

-إذا كان حجم العينة كبير أكبر أو يساوي 30 $(n \geq 30)$ نستخدم اختبار التوزيع الطبيعي Z

- إذا كان حجم العينة صغير أصغر أو يساوي ($30n <$) نستخدم اختبار توزيع T ستودنت.

-خطوات اختبار الفرض الإحصائي:

1- صياغة الفرض الصفري والفرض البديل

2-تحديد حجم العينة، متوسط العينة و الانحراف المعياري.

3-نحدد عبارة الاختبار الإحصائي($z = \frac{x-\mu}{\sigma/\sqrt{n}}$)

Xيمثل متوسط العينة، μ متوسط المجتمع، σ الانحراف المعياري، n حجم العينة.

4-التطبيق العددي للاختبار الإحصائي.

5-اتخاذ القرار بقبول أو رفض الفرض الإحصائي.

تذكير بالقيم الجدولية للاختبار الزاديZ :

مستوى الدلالة	**اختبار أحادي الطرف**	**اختبار ثنائي الطرف**
0.05	1.65	1.96
0.01	2.33	2.58

مثال1: إذا كان متوسط أعمار عمال إحدى المؤسسات سنة 2010 هو 36 سنة وفي عام 2022 أخذت عينة مكونة من 64 فردا من عمال المؤسسة، فوجد أن المتوسط الحسابي لأعمارهم هو 40 سنة والانحراف المعياري 8 سنوات، هل يدل هذا على أن متوسط أعمال العمال سنة 2010 قد اختلف عن متوسط أعمارهم سنة 2022 وذلك عند مستوى دلالة 0.05.

الحل:

S=8 x=40 n=64

1-صياغة الفرضيات:

$H_0 : \mu=36$

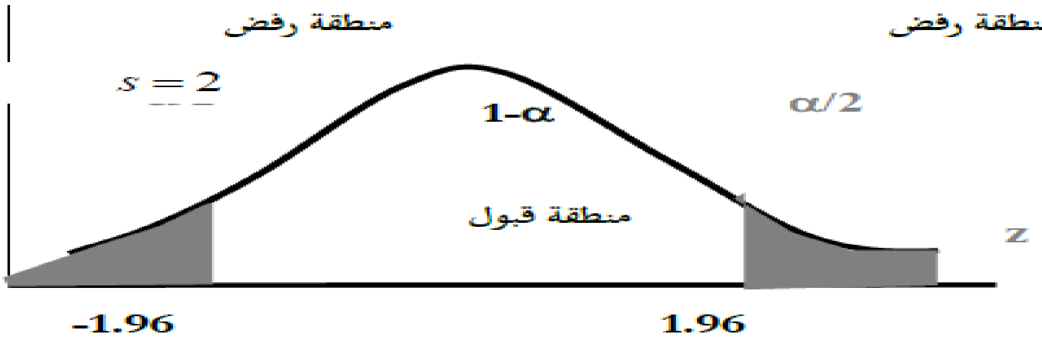
$H_1 : \mu\neq36$

2- بما أن $(n \geq 30)$ نستخدم اختبار التوزيع الطبيعي Z

($z = \frac{x-\mu}{\sigma/\sqrt{n}}$)

3-التطبيق العددي: $z = \frac{40-36}{\frac{8}{\sqrt{64}}}$

Z=4 المحسوبة



4-نلاحظ أن Z المحسوبة تقع في منطقة القبول وعليه نقبل الفرض الصفري ونرفض الفرض البديل، بمعنى أن متوسط أعمار العاملين سنة 2010لا يختلف عن متوسط أعمارهم سنة 2022.

مثال 2: اختيرت عينة من 64 تلميذ من إحدى المدارس المتوسطة فوجد أن متوسط أعمارهم 12 سنة، والانحراف المعياري سنتان(2سنة)،

المطلوب: هل يدل هذا على أن متوسط أعمار التلاميذ أكبر من 11 سنة(عند مستوى معنوية 0.01)

الحل: S=2 x=12 n=64

-صياغة الفرضيات:

$H_0 : \mu=11$

$H_1 : \mu>11$

2- بما أن $(n \geq 30)$ نستخدم اختبار التوزيع الطبيعي Z

$(z = \frac{x-\mu}{\sigma/\sqrt{n}})$

3-التطبيق العددي: $z = \frac{12-11}{\frac{2}{\sqrt{64}}}$

Z=4 المحسوبة

- نلاحظ أن Z المحسوبة تقع في **منطقة الرفض** وعليه نرفض الفرض الصفري و**نقبل الفرض البديل** أي أن متوسط أعمار التلاميذ بهذه المتوسطة أكبر من 11 سنة.

-اذا كان حجم العينة أقل من 30 فإننا نستخدم اختبار ت ستودنت. بدلا من اختبار زاد

مثال: إذا كانت أسعار إحدى السلع تتبع توزيع طبيعي، وكان متوسط هذه السلعة سنة 2007 هو 38 دج وفي عام 2018 اختيرت عينة من 16 وحدة من السلعة،فكان المتوسط الحسابي للسلعة 40 دج و بانحراف معياري 4دج.

المطلوب: هل تدل هذه البيانات على اختلاف متوسط السلعة في العامين ؟ استخدم مستوى معنوية 0.01

الحل:

الحل: S=4 x=40 n=16

-صياغة الفرضيات:

$H_0 : \mu=38$

$H_1 : \mu\neq38$

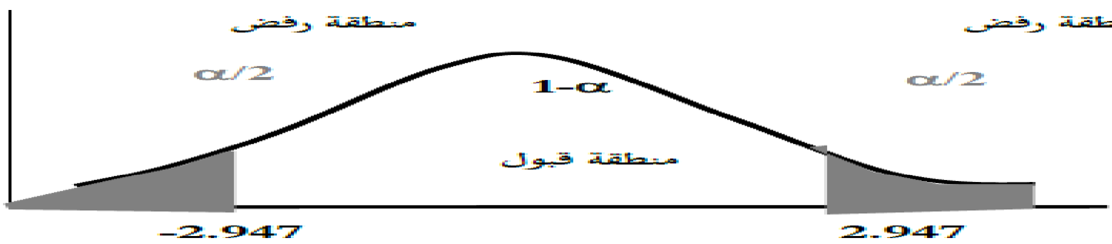
2- بما أن $(n < 30)$ نستخدم اختبار ت ستودنت

$(t = \frac{x-\mu}{\sigma/\sqrt{n}})$

3-التطبيق العددي: $t = \frac{40-38}{\frac{4}{\sqrt{64}}}$

t=2 المحسوبة

t (n-1,α/2) الجدولية ومنه t (15,0.005) =2.947



T المحسوبة تقع في منطقة القبول وعليه نقبل الفرض الصفري ونرفض الفرض البديل بمعنى أنه لا يوجد اختلاف في سعر السلعة بين العامين

ملاحظة: جميع الاختبارات يتم حلها بنفس المراحل(الطريقة) الاختلاف في المعطيات و قانون الاختبار الإحصائي.

-اختبار فرض حول نسبة في المجتمع:

نستخدم نفس خطوات السابقة ونطبق العلاقة التالية

$z = \frac{r-P}{\sqrt{\frac{p(1-p)}{n}}}$:

مثال: من 900 شخص وجد أن عدد المؤيدين منهم لرأي معين هو738 شخص

المطلوب: اختبر الفرض أن نسبة المؤيدين لذلك الرأي في المجتمع هو 0.8 ؟ عند مستوى معنوية 0.01

الحل: $r = 0.82$ $r = \frac{738}{900}$ n=900

-صياغة الفرضيات:

$H_0 : P=0.8$

$H_1 : P\neq 0.8$

$$z = \frac{r-P}{\sqrt{\frac{p(1-p)}{n}}}$$

-التطبيق العددي: $z = \frac{0.82-0.8}{\sqrt{\frac{0.8(1-0.8)}{900}}}$

Z=2 المحسوبة

نلاحظ أن Z المحسوبة تقع في منطقة القبول وعليه نقبل H_0 ونرفض H_1 أي أن عدد المؤيدين لهذا الرأي في المجتمع يساوي 0.8

-اختبار ذ-z للعينات الكبيرة(متساوية أو غير متساوية العدد)

عندما يكون حجم العينتين أكبر من 30 فان نسبة الفرق بين متوسطي العينتين منسوبا إلى الخطأ المعياري ويتم حسابها باستخدام نسبة **z** حيث يتم تفسير الدلالة الإحصائية بالرجوع إلى جداول الاحتمال لz من خلال المعادلة التالية:

$$z = \frac{(x1 - x2) - (u1 - u2)}{\sqrt{\left(\frac{S1^2}{n1}\right) + \left(\frac{S2^2}{n2}\right)}}$$

تتأسس هذه المعادلة على أن $\mu_1 - \mu_2 = 0$

حيث:

x_1 المتوسط الحسابي للعينة الأولى.

x_2 المتوسط الحسابي للعينة الثانية.

μ_1 متوسط المجتمع الذي سحبت منه العينة الأولى.

μ_2 متوسط المجتمع الذي سحبت منه العينة الثانية .

S_1^2 تباين العينة الأولى

S_2^2 تباين العينة الثانية.

n_1 عدد أفراد العينة الأولى

n_2 عدد أفراد العينة الثانية.

مثال: أخذت عينتان مستقلتان بغرض مقارنة متوسطات المجتمعين الإحصائيين الذي سحبت منهما العينتان فكانت البيانات على النحو التالي:

العينة الأولى: $n_1=50$ $x_1=57.5$ $S_1=6.2$

العينة الثانية: $x_2=54.4$ $n_2=60$ $S_2=10.6$

المطلوب: اختبر الفرض الذي يقرر أن متوسط المجتمع الذي سحبت منه العينة الأولى أكبر من المتوسط الذي سحبت منه العينة الثانية عند مستوى معنوية 0.04

الحل:

$H_0 : \mu_2 = \mu_1$

$H_1 : \mu_1 > \mu_2$

نقوم بتطبيق المعطيات على المعادلة التالية:

$$z = \frac{(x1 - x2) - (u1 - u2)}{\sqrt{\left(\frac{S1^2}{n1}\right) + \left(\frac{S2^2}{n2}\right)}}$$

نحصل على النتيجة التالية Z المحسوبة تساوي 1.90 نقارنها مع الجدولية 2.33 من خلال التمثيل البياني

نلاحظ أن Z المحسوبة(1.90) تقع في منطقة القبول وبالتالي نقبل الفرض الصفري ونرفض الفرض البديل وعليه فإنه لا توجد فروق بين متوسط المجتمع الذي سحبت منه العينة الأولى يساوي متوسط المجتمع الذي سحبت منه العينة الثانية.

محاضرة 03: اختبار T لعينتين مستقلتين / واختبار T لعينتين مرتبطتين

يتم التمييز بين عينتين من حيث الارتباط والاستقلال على أنه في حالة الارتباط يتم اختبار نتائج نفس العينة في اختبارين مختلفين قبلي وبعدي، أما في حالة الاستقلال يتم مقارنة نتائج العينتين في نفس الاختبار وقبل تطبيق الاختبار يجب التأكد من

- **الاستقلال**: لا يحتاج لاختبار إحصائي.
- **التجانس**: يتم التأكد من أن تباين العينة الأولى يساوي تباين العينة الثانية وللتأكد من التجانس يتم تطبيق اختبار يسبق اختبار ت ستودنت هو اختبار التجانس الفرض الصفري ينص أن تباين العينة الأولى يساوي تباين العينة الثانية ، بينما في حالة عدم التجانس فإن تباين العينة الأولى لا يساوي تباين العينة الثانية

$H_0 : S_1^2 = S_2^2$

$H_1 : S_1^2 \neq S_2^2$

-إذا تحقق التجانس نستخدم اختبار ت ستودنت

-إذا لم يتحقق التجانس نستخدم اختبار آخر مشابه يسمى اختبار فيشر .F-test

1/اختبار T لعينتين مستقلتين

نستخدم الفرضيات على النحو التالي:

$H_0: \bar{x}_1 = \bar{x}_2$

$H_1: \bar{x}_1 \neq \bar{x}_2$

الفرضية الصفرية: لا توجد فروق بين العينتين

الفرضية البديلة: توجد فروق بين العينتين.

$\bar{x}_1$: المتوسط الحسابي للعينة الأولى

$\bar{x}_2$: المتوسط الحسابي للعينة الثانية.

قانون الاختبار الإحصائي:

$$T = \frac{\bar{x}_1 - \bar{x}_2}{sp\sqrt{\frac{1}{n1} + \frac{1}{n2}}}$$

$$S^2_p = \frac{(n1-1)s_1{}^2 + (n2-1)s_2{}^2}{n1+n2-2}$$

مثال: أراد أحد الباحثين دراسة تأثير استخدام الأساليب الحديثة في التدريب الرياضي على تعلم المهارات في كرة القدم حيث طبق هذه الأساليب على عينة من الرياضيين الذين لم يسبق لهم التدرب على مثل هذه الأساليب وقام بمقارنة النتائج مع عينة من الرياضيين الذين استخدموا الأساليب التقليدية وكانت النتائج مدونة في الجدول التالي:

5	6	8	7	6	10	7	6	5	10	المجموعة1
2	3	6	5	4	8	7	5	3	7	المجموعة2

المطلوب: هل متوسط المجموعة الأولى أكبر من متوسط المجموعة الثانية ؟ عند مستوى معنوية 0.05

الحل:

صياغة الفرضيات:

$H_0: \bar{X}_1 = \bar{X}_2$

$H_1: \bar{X}_1 > \bar{X}_2$

نحتاج إلى معرفة درجة الحرية والتي تساوي $n1 + n2 - 2$

نحسب المتوسط الحسابي لكل من المجموعتين والذي يساوي مجموع القيم على عددها:

$\bar{X}_1=\frac{\sum x}{n}$=10+5+6+7+10+6+7+8+6+5/10=7

وبنفس الطريقة نحسب المتوسط الحسابي للمجموعة الثانية والذي نجده يساوي5

نطبق العلاقة السابقة:

$$T = \frac{\mathrm{x}\,1 - \mathrm{x}\,2}{sp\sqrt{\frac{1}{n1} - \frac{1}{n2}}}$$

نقوم بحساب التباين للعينة الأولى

$$s_{1^2} = \frac{\sum \mathrm{x}^2}{\mathrm{n}} - x_{1^2}$$

$$s_{1^2} = \frac{520}{10} - 49$$

ومنه تباين العينة الأولى يساوي 3

بنفس الطريقة نحسب تباين العينة الثانية:

$$s_{2^2} = \frac{\sum \mathrm{x}^2}{\mathrm{n}} - x_{1^2}$$

$$s_{1^2} = \frac{286}{10} - 25$$

ومنه تباين العين الثانية يساوي 3.6

وعليه نجد ان

$$S^2_p = \frac{(9)3 + (3.6)9}{20-2}$$

وعليه S^2_p تساوي 3.3

الان نطبق المعادلة الكلية لحساب قيمة ت ستودنت:

$$T = \frac{7 - 5}{\sqrt{3.3}\sqrt{\frac{1}{10} - \frac{1}{10}}}$$

وبعد العملية الحسابية نجد أن قيمة ت $-T$ المحسوبة تساوي 2.47

نقوم بالمقارنة بين ت $-T$ المحسوبة مع ت $-T$ الجدولية حيث1.73= T(20-2,0.05)

بما أن ت $-T$ المحسوبة أكبر من ت $-T$ الجدولية فإننا نرفض الفرض الصفري ونقبل الفرض البديل بمعنى أن البرامج التدريبية التي تستخدم الأساليب الحديثة فعالة.

2/اختبار T لعينتين غير مستقلتين(مرتبطتين)

يستخدم هذا الاختبار عندما يكون لدينا بيانات عينتين غير مستقلتين معنى ذلك عينة واحدة وكل مفردة في العينة تحمل قراءتين (قبلية وبعدية) القراءة الأولى تعتبر عينة أولى والقراءة الثانية نعتبرها عينة ثانية وفي هذه الحالة يكون الهدف هو اختبار الفرق بين بيانات الاختبار الأول والثاني **(d)** معنوية أم لا حيث d يمثل الفرق بين القياس القبلي والبعدي ؟وتصاغ فرضيات هذا الاختبار على النحو التالي:

لا توجد فروق بين العينتين H_0

توجد فروق بين العينتين H_1

$H_0 : d = 0$

$H_1 : d \neq 0$

وفي هذا الاختبار درجة الحرية تساوي n-1

وقانون هذا الاختبار يحسب من خلال المعادلة التالية:

$$T = \frac{\sum d}{\sqrt{\frac{n\sum d^2 - (\sum d)^2}{n-1}}}$$

$\sum d$: مجموع الفروق بين درجات القياس القبلي والبعدي

$\sum d^2$: مجموع مربعات الفروق بين درجات القياس القبلي والبعدي

$(\sum d)^2$: مربع مجموع الفروق بين القياس القبلي والبعدي

مثال: الدرجات التالية لقياسين قبلي وبعدي لمجموعة واحدة تتكون من 10 أفراد

القياس القبلي	58	45	61	55	58	90	26	35	42	48
القياس البعدي	42	50	23	50	45	85	30	20	50	60

المطلوب: حساب دلالة الفرق بين متوسطي القياسين القبلي والبعدي عند مستوى معنوية 0.05 تأسيسا على أن كل قياس يختص بمجموعة واحدة وأن المجموعتين مرتبطتين، وأن الأفراد تعرضوا لمتغير تجريبي بعد القياس القبلي

الحل:نقوم بصياغة الفرضيات(الصفرية والبديلة)

$H_0: \bar{X}_1 = \bar{X}_2$

$H_1: \bar{X}_1 \neq \bar{X}_2$ فرض بديل غير موجه

$$T = \frac{\sum d}{\sqrt{\dfrac{n\sum d^2 - (\sum d)^2}{n-1}}}$$

لكي يتم تطبيق هذه المعادلة نلخص الحل في الجدول الموالي:

القياس القبلي	58	45	61	55	58	90	26	35	42	48	المجموع
القياس البعدي	42	50	23	50	45	85	30	20	50	60	
الفرق بين القياسين-d-	16	5-	38	5	13	5	4-	15	8-	12-	63
$\sum d^2$	256	25	1444	25	169	25	16	225	64	144	2393
	/										$=3969(\sum d)^2$

$$T = \frac{63}{\sqrt{\dfrac{23930 - 3969}{9}}}$$

المحسوبة $T = 1.33$

نقوم بإيجاد T الجدولية حيث درجة الحرية تساوي 9 ومستوى الدلالة 0.05 /2(اختبار ثنائي الطرف)

T(9,0.025) =2.26

القرار:نلاحظ أن قيمة ت الجدولية أكبر من ت المحسوبة وعليه نقبل الفرض الصفري ونرفض الفرض البديل أي انه لا توجد فروق ذات دلالة إحصائية بين القياسين القبلي والبعدي.

محاضرة 04: تحليل التباين ANOVA

في اختبار ت ستودنت للعينات المستقلة Independent T Test أنه يوجد متغير واحد بمستويين ومتغير تابع واحد، وكمثال عن ذلك عند المقارنة بين الذكور والإناث(الجنس متغير مستقل بمستويين هما ذكور وإناث) في اختبار مهاري في رياضة ما(متغير تابع بمستوى واحد هو نتيجة الاختبار المهاري)، ولكن عندما يكون المتغير المستقل بأكثر من مستويين فإننا نستخدم اختبار تحليل التباين.

يستخدم اختبار تحليل التباين لاختبار فرضية اختلاف المتوسطات الحسابية لعدد من المتغيرات المستقلة مثلا نقارن بين اللياقة البدنية بين لاعبي ثلاث رياضيات هي ألعاب القوى، كرة القدم، وكرة اليد، أو نقارن بين ساعات المذاكرة بين طلبة التخصصات الثلاث(تدريب رياضي، تربية حركية، إدارة رياضية، إعلام رياضي)، نلاحظ من خلال المثالين انه يوجد متغير مستقل واحد بثلاث مستويات أو أكثر ومتغير تابع واحد فنوع الرياضة متغير مستقل وله ثلاث مستويات وليس ثلاث متغيرات، وكذا التخصص الدراسي بأربع مستويات وليس أربع متغيرات ويسمى المتغير المستقل في تحليل التباين ب (Factor) والاختبار الذي فيه متغير مستقل واحد ومتغير تابع واحد كما ذكرنا في المثالين السابقين يسمى تحليل التباين الأحادي(One-WAY ANOVA).

وفي حالة إضافة متغير مستقل آخر للمثال الثاني الذي يتعلق بساعات المذاكرة وليكن متغير جنس الطالب(ذكور وإناث) يصبح لدينا متغير تابع واحد هو ساعات المذاكرة ومتغيرين مستقلين اثنين هما جنس الطالب ونوع التخصص الدراسي وفي هذه الحالة يسمى هذا الاختبار بالتباين الثنائي(TWO-WAY ANOVA) وإذا ما تابعنا الموضوع وأضفنا متغير مستقل آخر يصبح الاختبار يسمى تحليل التباين الثلاثي(THREE-WAY ANOVA) مثل دراسة تفاعل جنس الطالب ومستواه الدراسي وحالته الاجتماعية على التحصيل الدراسي وهنا يصبح لدينا ثلاث متغيرات مستقلة لكل منها مستويات وهذا الاختبار غير شائع كثيرا.

ولكن عندما يكون لدينا متغيرين تابعين اثنين ومتغير مستقل واحد مثل دراسة اللياقة البدنية والأداء الرياضي تبعا لنوع الرياضة فإن الاختبار الجديد يسمى اختبار التباين المتعدد(MANOVA) وفي هذه المحاضرة سوف نكتفي بشرح طريقة حساب اختبار تحليل التباين الأحادي والجدول الموالي يوضح الفرق بين أنواع اختبارات تحليل التباين:

الاختبار	عدد المتغيرات المستقلة	عدد المتغيرات التابعة
تحليل التباين الأحادي	متغير واحد=المستوى الدراسي	متغير واحد=ساعات المذاكرة
تحليل التباين الثنائي	متغيرين : الجنس+ المستوى الدراسي	متغير واحد=ساعات المذاكرة
تحليل التباين المتعدد	متغير واحد=المستوى الدراسي	أكثر من متغير=عدد ساعات المذاكرة+الذكاء

مثال: اراد احد الباحثين إجراء تجربة للتعرف على تأثير التدريب العقلي والتدريب البدني والتدريب البدني والتدريب العقلي معا على تعلم مهارة حركية ما فقام باختيار 4 مجموعات تتكون كل واحدة من 6 أفراد ثم قام بتوزيع المجموعات عشوائيا على برامج التدريب التي استمرت لمدة 4 أسابيع

بمعدل 30 دقيقة يوميا للمجموعات التجريبية في حين لم تتدرب المجموعة الضابطة على أي برنامج وفي نهاية المدة قام الباحث باختبار المجموعات الأربع في المهارة الحركية وتحصل على النتائج التالية:

تدريب عقلي	تدريب عقلي وبدني	تدريب بدني	العينة الضابطة	المجموع
16	17.5	17.5	15	66
17	19	18.5	16.5	71
17.5	19.5	19.5	17	73.5
18.5	21	20.5	18	78
19	21	21.5	18	79.5
17.5	20	19.5	16	73
$\sum = 105.5$	$\sum d = 118$	$\sum = 117$	$\sum = 100.98$	$\sum = 441.48$
$\bar{X}$=17.58 المتوسط	$\bar{X}$=19.66	$\bar{X}$=19.50	$\bar{X}$=16.83	$\bar{X}$=18.39

المطلوب: هل توجد فروق ذات دلالة إحصائية بين المجموعات الأربع عند مستوى دلالة 0.05؟

الحل:

صياغة الفرضيات

$H_0: \bar{X}_1 = \bar{X}_2 = X_3$

$H_1: \bar{X}_1 \neq \bar{X}_2 \neq X3$

مصدر التباين	مج المربعات	درجة الحرية	متوسط المربعات	F
بين المجموعات	SSA=r Σ $(x_j\text{-}x)^2$	c-1	MSSA=$\frac{SSA}{c-1}$	F=$\frac{\mathrm{MSSA}}{\mathrm{MSSE}}$ F=7.71
داخل المجموعات	SSE=SST-SSA	(r-1)c	MSSE=$\frac{SSE}{(\mathrm{r}-1)\mathrm{c}}$	
المجموع الكلي	SST= Σ $(x_{ij}\text{-}x)^2$	r.c-1		

SSA	0.65+1.61+1.23+2.43 =5.92*6=35.54	3	11.87	
SSE	66.49-35.62=30.87	20	1.54	
SST	7.71+….=66.49	23		

نلاحظ أن قيمة اختبار ف –F المحسوبة تساوي7.71

علينا مقارنة نتيجة الاختبار المحسوبة نقوم بإيجاد ف –F الجدولية(التقاطع بين القيمة الأفقية لدرجة الحرية بين المجموعات والقيمة العمودية لدرجة الحرية داخل المجموعات ومستوى الدلالة)**(F,3,20,0.05)=3.10**

بما أن قيمة ف –F المحسوبة أكبر من ف –F الجدولية نرفض الفرض الصفري ونقبل الفرض البديل وعليه توجد فروق ذات دلالة إحصائية بين المجموعات الأربع.

طريقة ثانية

يمكن استخراج تحليل التباين(القيمة الفائية) وفق المعادلة الثانية:

$$\text{ف}(F)=\frac{\text{متوسط المربعات بين المجموعات}}{\text{متوسط المربعات داخل المجموعات}}$$

$$\text{متوسط المربعات بين المجموعات}=\frac{\text{مجموع المربعات بين المجموعات}}{\text{درجات الحرية بين المجموعات}}$$

$$\text{متوسط المربعات داخل المجموعات}=\frac{\text{مجموع المربعات داخل المجموعات}}{\text{درجات الحرية داخل المجموعات}}$$

درجات الحرية بين المجموعات= عدد المجاميع-1

درجات الحرية داخل المجموعات=حجم العينة الكلي-عدد المجاميع

مثال: يمثل الجدول الموالي نتائج اختبار السحب على العقل للسنوات الثلاث في كلية التربية البدنية والرياضية لقياس مطاولة القوة العضلية المتحركة الذراعين:

أ	ب	ج	أ^2	ب^2	ج^2
5	7	3	25	49	9
2	7	4	4	49	16
6	7	7	36	49	49
5	8	3	25	64	9
4	9	2	16	81	4
مج=22	**مج=38**	**مج=19**	**مج=106**	**مج=292**	**مج=87**

نقوم بإيجاد معامل التصحيح(ح)

ح=(مج س)^2/ن

ح=(5+2+6+.......+3+7+2)^2/15

ح=(79)^2/17

ح=416.07

- إيجاد مجموع المربعات الكلي=($5^2+2^2+6^2+....+3^2+7^2+2^2$)-ح

مجموع المربعات الكلي=416.07-485

- إيجاد مجموع المربعات بين المجموعات=(22/5)^2+(38/5)^2+(19/5)^2-ح

مجموع المربعات بين المجموعات=416.07-457.8

- إيجاد مجموع المربعات داخل المجموعات=مجموع المربعات الكلي-مجموع المربعات بين المجموعات

مجموع المربعات داخل المجموعات=41.73-68.94

- متوسط المربعات بين المجموعات=2/41.73
- متوسط المربعات داخل المجموعات=12/27.2
- ف-F=2.26/20.86

ف-F=9.23 المحسوبة

محاضرة 5: اختبار كاف -تربيع- Chi-Square

اختبار كاف- تربيع من أشهر الاختبارات اللامعلمية، وهو مبني على مقارنة التكرارات المشاهدة مع التكرارات المتوقعة لمعرفة هل توجد فروق معنوية بينهما أم لا؟

شروط اختبار كا^2-X^2:

***عشوائية العينة.**

***استقلال المشاهدات.**

***حجم العينة أكبر من 30.**

يستخدم هذا الاختبار في الحالات التالية:

- **جودة التوفيق.**
- **الاستقلال.**
- **التجانس.**

1-اختبار جودة التوفيق:

يتم تطبيقه في حالة استخدام أسئلة الاستبيان لمعرفة دلالة الفرق لصالح القيمة الاكثر تكرار من عدمه وتتم صياغة الفرضيات على النحو التالي:

-الفرض الصفري(H_0): لا توجد فروق ذات دلالة إحصائية لصالح القيمة الأكثر تكرار.

-الفرض البديل(H_1): توجد فروق ذات دلالة إحصائية لصالح القيمة الأكثر تكرار.

عبارة الاختبار الإحصائي هي كما يلي:

$$\mathbf{x^2} = \sum \frac{(f_o - f_e)2}{f_e}$$

حيث:

f_o: التكرار المشاهد(الحقيقي)

f_e : التكرار المتوقع.

$$f_e = \frac{n}{k}$$

n : مجموع التكرار(حجم العينة)

k : عدد البدائل

-درجة الحريةdf: تساوي (عدد البدائل-1)وبعبارة أخرى(df=k-1)

مثال:الجدول التالي يوضح إجابات أفراد على أحد أسئلة الاستبيان

البدائل	التكرار المشاهد(الحقيقي)f_o	التكرار المتوقعf_e
غير موافق بشدة	8	10
غير موافق	5	10
أحيانا	7	10
موافق	5	10
موافق بشدة	25	10
المجموع	50	10

المطلوب: هل توجد فروق ذات دلالة إحصائية لصالح القيمة الأكثر تكرار عند مستوى دلالة0.05؟

الحل:

1-صياغة الفرضيات:

-الفرض الصفري(H_0): لا توجد فروق ذات دلالة إحصائية لصالح القيمة الأكثر تكرار.

-الفرض البديل(H_1): توجد فروق ذات دلالة إحصائية لصالح القيمة الأكثر تكرار.

2-نقوم بتطبيق العلاقة :

$$x^2 = \sum \frac{(f_o\_f_e)2}{f_e}$$

3-التطبيق العددي:

$$x^2 = \frac{(8-10)^2 \quad + (5-10)^2 \quad (7-10)^2 \quad (5-10)^2 \quad (25-10)^2}{10}$$

$$x^2 = \frac{4+25+9+25+225}{10}$$

$$\text{المحسوبة } x^2 = 28.80$$

4-نقارن القيمة المحسوبة مع القيمة الجدولية: df: تساوي (عدد البدائل-1)وبعبارة أخرى(df=k-1)نجدها تساوي4

X^2 (0.05,4)=9.488

5-القرار: نلاحظ أن قيمة كا^2 المحسوبة أكبر من قيمة كا^2 الجدولية وعليه نرفض الفرض الصفري ونقبل الفرض البديل بمعنى أنه توجد فروق ذات دلالة إحصائية لصالح القيمة الأكثر تكرار(موافق بشدة=25)

2-اختبار الاستقلال:

يستخدم في حالة معرفة إذا ما كان هناك علاقة بين صفتين من صفات المجتمع أم لا.

مثال: نريد معرفة العلاقة بين مستوى الدخل ومستوى التعليم، أو العلاقة بين التدخين ومستوى التعليم، أو العلاقة بين الوزن ولون العينين في مجتمع ما .

الفرضيات:

-الفرض الصفري(H_0): لا توجد علاقة بين الصفتين.

-الفرض البديل(H_1): توجد علاقة بين الصفتين.

-نقوم بحساب التكرار المتوقع في كل خانة(خلية)

$$f_e = \frac{\text{مجموع العمود الذي به الخلية} X \text{مجموع الصف الذي به الخلية}}{\text{المجموع الكلي}}$$

2-نقوم بتطبيق العلاقة :

$$x^2 = \sum \frac{(f_o\_f_e)2}{f_e}$$

-درجة الحرية **df : تحسب كما يلي:(عدد الأعمدة-1)(عدد الصفوف-1) =df**

مثال تطبيقي:

لدراسة العلاقة بين التعليم والتدخين تم اختيار عينة عشوائية حجمها 400 شخص والنتائج مدونة في الجدول التالي:

التعليم	**التدخين**		المجموع
	يدخن	**لا يدخن**	
بدون تعليم	170	50	220
تعليم متوسط	50	70	120
تعليم عالي	20	40	60
المجموع	240	160	400

المطلوب: هل توجد علاقة بين التدخين ومستوى التعليم عند مستوى دلالة 0.05؟

الحل:

1-صياغة الفرضيات:

-الفرض الصفري(H_0): لا توجد علاقة بين الصفتين.

-الفرض البديل(H_1): توجد علاقة بين الصفتين.

2-نقوم بحساب التكرارات المتوقعة لكل خانة(خلية):

$$f_{e1} = \frac{220X\,240}{400}$$

$$f_{e1} = 132$$

$$f_{e2} = \frac{120X\,240}{400}$$

$$f_{e2} = 72$$

وبنفس الطريقة نحسب باقي التكرارات المتوقعة

-3-نقوم بتطبيق العلاقة :

$$x^2 = \sum \frac{(f_o \_ f_e)2}{f_e}$$

4-التطبيق العددي:

$$x^2 = \frac{(170-132)^2}{132} + \frac{(50-88)^2}{88} + \cdots$$

$$x^2 = 10.94 + 16.41 + \cdots$$

في مثل هذه الحالات لا داعي لمواصلة الحساب لأن قيمة x^2 الجدولية هي اصغر من المحسوبة

-x^2 الجدولية ؟ درجة الحرية:**(عدد الأعمدة-1)(عدد الصفوف-1) =df نجد، df = 2**

$X^2(0.05,2)=5.991$

5-القرار: بما أن كا^2 المحسوبة أكبر من كا^2 الجدولية نرفض الفرض الصفري ونقبل الفرض البديل الذي يقرر أنه توجد علاقة ذات دلالة إحصائية بين التدخين ومستوى التعليم.

محاضرة 6: الارتباط والانحدار الخطي

تعريف الارتباط:

هو علاقة بين متغيرين(ظاهرتين) (X,Y) بحيث انه إذا تغير احد المتغيرين يتبعه المتغير الآخر في نفس الاتجاه فيكون الارتباط طردي مثل العلاقة الدخل والاستهلاك، وقد يكون الاتجاه متعاكس فيكون الارتباط عكسي مثل العلاقة بين الاستهلاك والادخار، أما في حالة استقلال الظاهرتين فإن الارتباط يكون منعدم مثل العلاقة بين الطول والذكاء.

1-معامل الارتباط بيرسون(الخطي):

هو معامل رقمي يوضح نوع ودرجة العلاقة بين متغيرين ويرمز له بالرمز(r)

مثال: تم اختيار عينة من الطلبة (n) وتم حساب أوزانهم(x) و أطوالهم(y)

الأوزان(x):$x_1,x_2,x_3...x_n$

الأطوال(y):$y_1,y_2,y_3,...y_n$

ويحسب معامل الارتباط بيرسون بالعلاقة التالية:

$$r = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{s_x \cdot s_y}$$

حيث:

$\bar{X}$ يمثل المتوسط الحسابي للمتغير X ويحسب بالعلاقة: $x = \frac{\sum x}{n}$

$\bar{y}$ يمثل المتوسط الحسابي للمتغير y ويحسب بالعلاقة: $y = \frac{\sum y}{n}$

S_x: يمثل الانحراف المعياري للمتغير X يحسب : $s_x = \sqrt{\frac{\sum x^2}{n} - \bar{x}^2}$

S_y يمثل الانحراف المعياري y ويحسب بالعلاقة: $s_y = \sqrt{\frac{\sum y^2}{n} - \bar{y}^2}$

-**خصائص معامل الارتباط:**

- **إشارة r تكون موجبة في الارتباط الطردي وسالبة في الارتباط العكسي ، وتساوي صفر في حالة الارتباط المنعدم.**
- **قيمة r تساوي 1+ في حالة الارتباط الطردي التام، وتساوي-1 في حالة الارتباط العكسي التام.**
- **قيمة r تكون في المجال-1<r<+1 وتزداد قوتها كلما اقتربت من الواحد الصحيح.**

مثال1: من أجل دراسة العلاقة بين الدخل (X)والاستهلاك (Y)بآلاف الدولارات كانت النتائج التالية:

$$\sum x = 100 / n = 40 / \sum y^2 = 410 / \sum x^2 = 720 / \sum xy = 516 / \sum y = 120$$

المطلوب: حساب معامل الارتباط بيرسون بين الظاهرتين.

لتطبيق معادلة الارتباط التالية:

$$r = \frac{\frac{\sum xy}{n} - \bar{x}\,\bar{y}}{s_x \cdot s_y}$$

1-نقوم بحساب المتوسط الحسابي للظاهرة **X** والمتوسط الحسابي للظاهرة Y بتطبيق المعادلتين:

$x = \frac{\sum x}{n} / y = \frac{\sum y}{n}$ و بالتطبيق العددي $y = \frac{100}{40} / x = \frac{120}{40}$

وعليه نجد $\bar{X}=3$، $\bar{y}=2.5$

2-الآن نقوم بحساب الانحراف المعياري للظاهرتين S_x وS_y بتطبيق المعادلتين السابقتين فنجد:

$S_x=3$، $S_y=2$

3-الآن نقوم بالتطبيق العدد لمعادلة الارتباط بعدما صارت كل المعطيات متوفرة فنجد أن $r=0.9$ وهو ارتباط طردي قوي.

مثال2: الجدول الموالي يوضح نتائج مجموعة من الطلبة لمعرفة العلاقة بين مؤشر كتلة الجسمBMI (كلغ/م^2) وسباق السرعة(ثا) :

مؤشر كتلة الجسم(كلغ/م^2)	15	33	19	40	35	16	25	20
سباق السرعة(ثا)	10	20	13	30	25	11	18	14

-**المطلوب** : حساب معامل الارتباط بيرسون بين الظاهرتين.

2-معامل الارتباط سبيرمان(للرتب):

يستخدم معامل الارتباط سبيرمان في حالة البيانات الوصفية وكذا في حالة الأعداد الكبيرة، ويعطي قيمة تقريبية فقط لمعامل الارتباط بين الظاهرتين، ونحصل عليه بترتيب البيانات تصاعديا او تنازليا ثم نحسب الارتباط بين رتب الظاهرتين بدلا عن قيمهما باستخدام المعادلة التالية:

$$r = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث :d يمثل الفرق بين رتب الظاهرة (x) والظاهرة (y)

مثال: أحسب معامل الارتباط سبيرمان بين تقديرات مجموعة من الطلبة في مقياسي الإحصاء والثقافة البدنية من البيانات التالية:

مقياس الإحصاء(x)	D	C	B	C	D	C	A	F
مقياس الثقافة البدنية(y)	F	D	A	D	F	B	C	D