الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة الشهيد حمة لخضر الوادي

قسم العلوم الاجتماعية

محاضرات في مقياس المعالجة الإحصائية للبيانات

محاضرة رقم 03 مدخل إلى الاحصاء المعمق

موجهة لطلبة السنة أولى ماستر إرشاد وتوجيه

إعداد الأستـــــــاذ:

د. محمد سبــع

السنة الجامعية: 2020 / 2021

محتوى المادة:

3. مدخل إلى الاحصاء الاستدلالي

المعلم (parametric) مفردة تعني صفة أو خاصية لمجتمع معين في مقابل تقدير (estimate) التي تكون صفة أو خاصية لعينة ما ، وأهم ما يميز الاحصاء المعلمي عن اللامعلمي هو الوسط الحسابي والانحراف المعياري ولذلك فان علم الاحصاء يميز بين شروط اختبار (ت) للعينات المستقلة و (ت) للعينات المترابطة لأن الأخير لا يتعامل مع أوساط وانحرافات، بينما نلاحظ أن اختبار (ت) للعينات المستقلة يتعامل معها أسوة بقوانين تحليل التباين (F) وقوانين (Z) ، من هنا يمكن فهم الاحصاء المعلمي أنه مجموعة من الطرق التي تتطلب تحقق افتراضات محددة حول المجتمع الذي تسحب منه العينة –وهنا مقتضى الدقة الانتباه للتعبير (حول المجتمع) كونه يختلف عن (حول العينة)، وبالتالي فإن الاحصاء اللامعلمي هو مجموعة من الطرق البديلة التي تستخدم في حالات عدم تحقق الافتراضات حول المجتمع الذي تسحب منه العينة أو في حالة البيانات الإسمية والرتبية، وكلا الإحصائين (المعلمي،اللامعلمي) من طرق الاحصاء الاستدلالي التي يمكن تعميم نتائجها على المجتمع إلا أن لكل منها مستوى ثقة معين يتحدد على ضوء البيانات المتوفرة وكذلك شروط تحقق الافتراضات .

1.3.شروط الإحصاء البارامتري (المعلمي)

لا يختلف الاحصائيون على أن هناك مجموعة من الافتراضات أو الشروط التي يجب توافرها لكي نستطيع أن نتعامل مع البيانات بالطرق المعلمية* والتي باختلال أي منها يحصل عدم اطمئنان من النتائج المستخرجة بهذه الطرق مما يعني اللجوء الى طرق أخرى لمعالجتها ، وهذه الشروط هي :

.* يتعين على القارئ الانتباه هنا أننا أشرنا إلى شروط الإحصاء البارامتري (المعلمي) ولم نقل اللابارامتري لأن الأصل في اختيار أسلوب المعالجة الإحصائية هو أن يتبع الباحث الأساليب الاحصائية البارامترية التي تتطابق فيها خصائص العينة مع خصائص المجتمع الإحصائي وهو ما يمنح نتائج المعالجة الإحصائية موثوقية أعلى من تلك اللابرامترية التي تصبح بديلا للأولى في حالة عدم توفر الشروط الضامنة لاتباع أحد الاساليب البارامترية

أولا: التوزيع الطبيعي :

حسب نظرية النهاية المركزية فإنه كلما زاد عدد أفراد العينة كلما اقترب تباينها من تباين المجتمع ويمكن اعتبار أن التوزيع يكون طبيعيا بصورة تقريبية عندما يصبح حجم العينة (30) فما فوق.

ويعلق على هذا الشرط بكونه متعلق بقياس تباين العينة إلى تباين المجتمع، إذ وجد أنه كلما اقترب حجم العينة من (30) وصعوداً فإن تباينه سيقارب تباين حجم العينات الكبيرة (المئات والالاف) ومن هنا وضع الحد الفاصل (30) في التعامل مع بعض الوسائل الاحصائية من قبيل اختبار (ت) وتم التعامل معه كونه من المسلمات وعمم من خلاله فكرة أنه إذا كان عدد العينات أو المشاهدات أقل من (30) فإن شرط اعتدالية التوزيع (التوزيع الطبيعي) قد اختل وبالتالي وجب الانتقال إلى الإحصاء البديل (اللامعلمي) ، ويتفق العديد من الباحثين جزئيا مع هذه المقولة إذ أن شرط الاعتدالية للتوزيع يتحقق بالعدد (30) ويكون التوزيع طبيعياً ولكن يكون الكلام فيما لو قل العدد عن (30) حيث لا يعني بالضرورة فقدان هذا الشرط (التوزيع الطبيعي) إذ أن الأمر يكون خاضعا حينها لخصائص البيانات المأخوذة من العينات، ومن هنا نلاحظ تساهل البعض مع هذا الشرط في حدود الاعداد من (20-29)، لكننا نجد التشدد واضحا فيما لو قلت الأعداد عن (15) إذ ينصح الكثيرون باللجوء إلى البديل ، وهذا الكلام مدعوما بمعادلة الخطأ المعياري :

[14]

بمعنى أن التناسب بين الخطا المعياري وحجم العينة يكون عكسيا فكلما زاد حجم العينة كلما قل الخطأ المعياري وكلما قل حجم العينة (والحالة هذه) كلما زاد الخطأ المعياري أي أن الخطا إذا زاد فإن عملية تعميم النتائج لن تكون ممكنة.

$$\text{الخطا المعياري} = \frac{ع}{\sqrt{ن}}$$

أقول : إن القول السابق وإن كان ! ... من عدة وجوه :

أ. إن القول السابق معارض بقول مجموعة من كبار الاحصائيين العالميين من أمثال (Glass) و (Hopkins) اذ يرون اأن افتراض التوزيع الطبيعي يمكن مخالفته بدون تبعات تذكر.

ب. أن معادلة الخطأ المعياري تضع حجم العينة تحت الجذر بمعنى أنها تقلل من تأثيره إذ لا يخفى أن جذر أي عدد هو أقل من العدد نفسه ومثاله الفارق بين العدد (25) وجذره (5) وهنا نورد المثال التالي :

لو كان هناك عينتين الأولى حجمها (15) والثانية حجمها (30) ولكل منهما نفس الانحراف المعياري ولنفرض أنه 3 ، فاننا وبعد تطبيق المعادلة آنفة الذكر سيتبين أنه في الحالة الأولى (العينة 15) ستبلغ قيمة الخطأ المعياري (0.77) ، أما في الحالة الثانية (العينة 30) فانه ستبلغ قيمة الخطا المعياري (0.55) ومن ملاحظة القيمتين سنجد أن الفارق ليس بالشيء الكبير الذي يتوقف عليه تعميم النتائج .

ج . إن فلسفة افتراض التوزيع الطبيعي أي الحد (30) هو ليس السبب في حد ذاته إذ أنه من الجدير بالذكر الاشارة بأنه كلما قل حجم العينة ستظهر مشكلة من نوع آخر وهي مشكلة القيم الشاذة أو الالتواء الحاصل نتيجة عدم توزع البيانات طبيعيا لقلتها.

الا أن هذه المشكلة –وإن كانت قائمة في حالات معينة– فأنه يمكن تخطيها وذلك بالتأكد من عدم إلتواء التوزيع باختبار القيم احصائيا بوسائل متعددة من قبيل (مربع كاي ، سميرنوف –كولموجروف) أو بالكشف عن اعتدالية التوزيع من خلال رسم البيانات على منحنى كاوس أو بطريقة الساق والأوراق وكل هذه الطرق يمكن استخدامها في البرنامج الاحصائي الشهير (SPSS)، فإن ظهر أن التوزيع طبيعي أو غير ملتوي فإننا نكون حينها بمأمن من تبعات هذا الشرط وإن ظهر وجود التواء فليس أمامنا إلا استبعاد القيم المسببة للالتواء من العينة (إن أمكن ذلك) وإن لم يمكن ذلك فإننا سنكون مضطرين الى استخدام أحد الأساليب اللابارامترية.

د. يرى البعض أن شرط التوزيع الطبيعي هو للمجتمع وليس للعينة إذ يذكر (عبد الجبار توفيق) أنه يجب الإيفاء بافتراض التوزيع الطبيعي للمجتمع والذي هو غالباً ما يكون متحققاً .

وأخيرا ومن كل ما تقدم فاننا نرى بأن هذا الشرط لا يمثل عائقا كبيرا أمام استخدام الأساليب البارامترية (المعلمية) إذ يمكن تجاوزه مبدئياً أو إجرائياً باحدى الطرق التي تم ذكرها .

ثانيا: الاستقلالية

مفهوم الاستقلالية في مقابل مفهوم الارتباط، فإذا كان الارتباط يعني الاقتران في التغير بين متغيرين أو أن التباين في المتغير (س) يرافقه تباين في المتغير (ص) فإن الاستقلالية تعني أن قيمة الارتباط بين (س) و (ص) منعدمة أو تساوي صفرا عند استخدام عدد من العينات أو المشاهدات وهذا يقتضي أن يتم اختيار كل من العينتين عشوائيا من مجتمعاتها ومن هنا نلاحظ أن هذا الشرط لا يتوفر في حالة إجراء اختبار قبلي وبعدي لنفس العينة أي اختبار (ت) للعينات المترابطة لأنه يكون على عينة واحدة وليس عينتين الأمر الذي يكون فيه هذا الاختبار مستغني عن هذا الشرط (الاستقلالية)، ومثال شرط الاستقلالية عينتين عشوائيتين من مجتمعين مختلفين أو عينتين عشوائيتين من مجتمع واحد مع عدم الخلط بين أفراد المجموعتين لأنه سيؤدي إلى خطر انهيار الشرط والذي يمنع استخدام الاحصاء البارامتري (المعلمي) وعندها يتم اللجوء إلى الطرق اللامعلمية للاستدلال.

وهنا لابد من الإشارة إلى أن هذا الشرط متوافر غالباً إذ نجد أن أكثر الباحثين يلجؤون إلى استخدام العينات العشوائية (من دون قصد) وبالتالي هذا الشرط تحصيل حاصل في أغلب البحوث التربوية.

ثالثا: تجانس التباين

ويعني هذا الشرط أن لكل من العينتين تبايناً لا يختلف عن تباين العينة الثانية، وعدم الاختلاف هذا لا يعني بالضرورة التطابق في قيمة التباينين بل يعني أنه ليس بينهما فرق معنوي، ويجري اختبار تجانس التباينين باستخدام قانون (F) نسبة إلى Fisher:

$$F = \frac{S^2_{\text{Largest}}}{S^2_{\text{Smallest}}} \qquad [15]$$

حيث تشير قيمة S² largest إلى قيمة التباين الأكبر

وتشير قيمة S² smallest إلى قيمة التباين الأصغر

إذ نستخرج القيمة المحسوبة لنقارنها بالقيمة الجدولية من جداول النسبة الفائية في الملاحق لغرض الحكم بتجانس التباين في حالة صغر القيمة المحسوبة عن الجدولية والعكس بالعكس ، وهنا نقترح التعامل بتسمية (تكافؤ التباين) كبديل عن الاسم المشهور لهذا الشرط (تجانس التباين) لأنه أنسب في المقام ، وممكن التعامل مع هذا الافتراض بعدة وجوه :

اولاً: اذا كانت العينتين متساويتين في عدد افرادهما أي أن (ن1=ن2) ، وهنا اتفق الاحصائيون بأنه يمكن مخالفة هذا الشرط وإهماله في حالة تساوي عدد أفراد العينتين ومثاله (إذا كان لدينا عينتين كل منهما تتكون من 20 فردا حيث ن1=20 و ن2=20) فهنا لاداعي للاستغراق في هذا الشرط وممكن غض النظر عنه ، ولعل السبب في ذلك يعود إلى أن الانحراف المعياري يعتمد بنسبة كبيرة على عدد العينة كون أن قانونه يحوي على قيمة (ن) في المقام وبتساوي قيمة ال(ن) تتساوى الانحرافات تقريباً:

$$\text{ع} = \sqrt{\frac{\text{مج}\,(\text{س} - \bar{\text{س}})^{٢}}{\text{ن} - ١}}$$

ثانياً: اذا كانت العينتين غير متساويتين بالعدد أي أن (ن1 ≠ ن2) بمعنى أن إحدى العينتين أكبر من الأخرى وفي هذه الحالة هناك احتمالان لا ثالث لهما :

أ.الاحتمال الاول : أن تكون العينة الكبيرة منتمية للمجتمع ذي التباين الكبير والعينة الصغيرة منتمية للمجتمع ذي التباين الصغير ، وفي هذه الحالة يمكن اغفال هذا الشرط إذ يكون الباحث في

وضع آمن ، والسبب في ذلك أن احتمال ارتكاب خطأ من النوع الاول (أن تكون الفرضية الصفرية صحيحة ويتم رفضها) يكون قليل إلى درجة يمكن إهماله.

ب.الاحتمال الثاني : أن تكون العينة الكبيرة منتمية للمجتمع ذي التباين الأقل والعينة الصغيرة منتمية للمجتمع ذي التباين الأكبر ، وهنا تكمن المشكلة إذا يكون هذا الشرط معرض للانتهاك نتيجة ارتفاع خطر ارتكاب خطأ من النوع الأول إلا أنه أيضاً لا يخلو من حالتين لا ثالث لهما :

الاولى : اذا قبلت الفرضية الصفرية : أي أن القيمة المحسوبة أصغر من القيمة الجدولية مما يعني أن الفروق مرجعها للصدفة وهي ليست معنوية وهنا لا نقع بالخطأ من النوع الأول ويكون الباحث بالجهة الامينة .

الثانية : إذا رفضت الفرضية الصفرية : أي أن تكون القيمة المحسوبة أكبر من الجدولية وبالتالي ليس للصدفة دور في الفروق وأن هناك فروق معنوية، وهنا يحدث التوقف غالباً، إذ نرى أن شرط تجانس التباين من أهم الشروط الثلاثة التي تم ذكرها وبالتالي فإنه لا يمكن التغافل عنه في الحالة المذكورة (الثانية) ، إلا أنه ومع خطر اختلال هذا الشرط فقد وضعت مجموعة حلول ممكن استخدامها لتجعلنا في الجانب الأمين من هذا الشرط وبالتالي يمكن إغفاله وهذه الحلول هي :

1.يمكن اجراء اختبار ليفين (Leven) وهو الاختبار المسؤول عن تجانس التباين والذي يحدد هل أن التباينات متجانسة أم لا ، فإذا ظهر عدم تجانسها فإننا نلجأ للخيار التالي.

2.تطبيق اختبار ويلتش (Welch) الذي يقوم بتعديل درجات الحرية، حيث ومن خلال هذه المعادلة يتم استخراج درجة حرية معدلة وهي غير درجة الحرية في الحالات الطبيعية وباستخراجها يتم استخراج القيمة الجدولية ليتم مقارنتها بالقيمة المحسوبة لـ (ت)

ومن هنا يمكننا رفض أو قبول الفرضية الصفرية فإذا رفضت لا يكون أمامنا الا اللجوء الى الاحصاء اللامعلمي إذ أنه حينها ستكون نتائجه أوثق من الطرق المعلمية.

2.3. معايير اختيار الاحصاء البارامتري (المعلمي) واللابارمتري (اللامعلمي)

تبين مما تقدم أن نوع البيانات وخصائص الأفراد والمشاهدات هو الحاكم الرئيسي في توجيه البحث نحو استخدام النوع المناسب (البارامتري أو اللابارمتري) حيث نرى أن البيانات تصنف على أربعة مقاييس وهذه المقاييس (مع الاختصار) هي:

أ.**المقاييس الاسمية** : مثل (ذكور-اناث) (عمال- موظفين) (أحمر-أصفر-....) وضابطه عملية التصنيف بمعنى التعامل مع أي متغير يمكن تصنيفه وهنا يعطى لكل صنف رقم لا يعبر إلا عن صنفه بحيث لا يمكن إجراء العمليات الحسابية الأربع عليه كأن يعطى للتصنيف (الذكور) رقم (1) و (الاناث) رقم (2) وهنا لا يمكن أن تجرى عملية جمع فنقول (1+2)لأنها أرقام معبرة عن أصناف وليست أعداد ، ويتم التعامل مع هذا النوع من المقاييس بالطرق اللابارمترية بغض النظر عن عدد أفراد العينة فيتبدل الوسط الحسابي بالمنوال والانحراف المعياري بالمدى و(ت) لعينة واحدة باختبار ذي الحدين ومربع كاي وسميرنوف كل بحسبه و(ت) المستقلة بمربع كاي وت مترابطة بماكنمار وتحليل التباين للعينات المترابطة بكوجران وللعينات المستقلة بمربع كاي وكذا الأمر بالنسبة لقوانين الارتباط حيث تستبدل بقوانين (فاي ومعامل الارتباط الرباعي وبايسيريال وبوينت بايسيريال ...) كل بحسبه .

ب.**المقاييس الرتبية**: وهو أفضل من سابقه إحصائيا وفيه ترمز الأعداد إلى رتب لتبين المواقع النسبية للاشياء وضابطه الترتيب من أعلى إلى أدنى مثل (أكبر – أصغر ..) أو (ممتاز-جيد جدا – جيد ...) أو (أثقل من – أخف من ..) إذ ترتب المشاهدات تصاعديا أو تنازليا ولا تمثل الأرقام فيها كميات بل رتب لذا لا تجرى عليها العمليات الأربع وأغلب استخدامه في استمارات الاستبيان (موافق- غير موافق) (موافق بدرجة عالية ...– موافق بدرجة منخفضة) إذ تعطى الدرجة (5) لـ (موافق بشدة) والدرجة (4) لـ (موافق بدرجة عالية) ...وهكذا ، وهنا يتم استخدام الاحصاء اللابارمتري فتستخدم اختبارات سبيرمان وكندال بدل ارتباط بيرسون وتستخدم قوانين ويلكوكسن ومان وتني بدل اختبارات (ت) المترابطة والمستقلة على التوالي وكذا يستخدم اختبار كوسكال واليز وفريدمان بدل تحليل التباين للعينات المستقلة والمترابطة على التوالي، وكذا يستخدم اختبار الوسيط وسميرنوف لعينتين مستقلتين على وفق تفصيلات سيأتي ذكرها لاحقا.

ج.**المقياس الفتري (الفاصل)**: وهو أعلى من سابقيه ويتميز بخاصية الفواصل والمسافات المتساوية التي تفصل بين درجة وأخرى ومن أمثلته (درجة الحرارة) و (مستوى الذكاء) إذ يختلف هذا المقياس عن سابقيه بأنه يمكن إجراء عمليات الجمع والطرح عليه وبالتالي إمكانية استخدام الأوساط الحسابية والانحرافات المعيارية وبالتالي استخدام الإحصاء البارامتري الذي يكون الأنسب معه وأهمها اختبار (ت) و (ف) ، وأهم ما يميز هذا المقياس أن الصفر فيه لايعتبر مطلقا أي أن الدرجة (صفر) لا تعني انعدام الصفة فعندما نقول إن درجة الحرارة هي (صفر) فهذا لا يعني انعدام درجة الحرارة بل إنها درجة كبقية

الدرجات لها دلالة معينة على خلاف ما إذا قلنا بأن طالبا ما في اختبار ما حصل (صفر) أي انعدام الإجابة وعدم تحقق أي شيء ، وتعتبر هذه الميزة (الصفر النسبي) هي ما يميز هذا المقياس عن المقياس النسبي.

وتجدر الاشارة الى أنه لا يمكن استخدام معامل الاختلاف في هذا المقياس وهي ميزة أخرى لهذا المقياس عن المقياس النسبي.

د.المقياس النسبي:

وهو أعلى مستويات القياس وأهمها وأكثرها ثقة من حيث النتائج الصادرة منه ويستخدم معه الإحصاء البارامتري (المعلمي) مثل (الطول –الوزن–المسافات–الزمن) ويكون فيه الصفر دال على انعدام الحالة فعندما نقول أن (س) من الأفراد حصل على صفر في مقياس ما فإنه يدل أن ليس لديه هذه الخاصية ويمكن إجراء جميع العمليات الحسابية عليه ويستخدم معه معامل الاختلاف أيضاً.

وبعد هذه الاشارة إلى مستويات القياس نقول أن المستويين الأولين (الاسمي والرتبي) يستخدم معها الطرق اللابارامترية (اللامعلمية) أما المقياسين (الفاصل والنسبي) فيستخدم معها الاحصاء البارامتري (المعلمي) إلا في حالات عدم توافر الافتراضات المذكورة سابقاً والتي تبين إمكانية التغلب على أكثر مشاكلها وبالتالي فلا يبقى هناك أي داع لاستخدام الأساليب اللامعلمية في المقاييس الفاصلة والنسبية إلا في حالات نادرة جدا، ومن هنا يتبين أن عملية العدول عن الإحصاء البارامتري المعلمي لمجرد عدم اكتمتال العدد (30) أو تجانس التباين أو الاستقلالية هو من الأخطاء الشائعة الاستخدام كما تم تفصيل شرحه سابقاً.

وهنا على الباحثين أن يضعوا في اعتبارهم أن الاختبارات البارامترية (المعلمية) هي أقوى وأدق في اختبار الفرضيات الإحصائية من الاختبارات اللابارمترية (اللامعلمية) وذلك لأنها تتحسس الفروق الموجودة في البيانات كونها تتعامل مع الانحرافات المعيارية وهذا ما يجعلها أعلى قدرة في ايجاد الفروقات وبالتالي فهي أكفأ وأقدر على رفض الفرضية الصفرية من الاختبارات اللابارمترية (اللامعلمية).

والجدول التالي يلخص بعض الأساليب الإحصائية ومستوى القياس الملائم.

جدول رقم (01) يوضح نوع المعالجة الإحصائية الملائم لنوع الفرضيات ونوع البيانات

عدد العينات	الفرض	التصميم التجريبى	نوع البيانات	الاختبار الإحصائى
عينة واحدة	التحقق من جودة المطابقة	مجموعة واحدة ذات الاختبار الواحد	اسمية	ذى الحدين – كا^2 –سمير نوف
			رتبية	سمير نوف – الإشارة
			فترية	– اختبار تZاختبار
عينتان مستقلتان	الفروق بين المجموعات	مجموعتان تجريبية – ضابطة	اسمية	كا^2 – فيشر – سمير نوف
			رتبية	الوسيط – مان ويتنى – التتابع
			فترية	اختبار ت
عينتان مترابطتان	الفروق بين القياسات	مجموعة واحدة ذات اختبارين قبلى وبعدى	اسمية	ماكنمار
			رتبية	ولكوكسن – الإشارة
			فترية	اختبار ت
عدة عينات مستقلة	الفروق بين المجموعات	المجموعات المتعددة	اسمية	كا^2
			رتبية	الوسيط– كروسكال ولاس
			فترية	تحليل التباين – تحليل التغاير
عدة عينات مترابطة	الفروق بين القياسات	مجموعة واحدة ذات الاختبارات المتعددة	اسمية	كوجران
			رتبية	فريدمان
			فترية	تحليل التباين ذى القياسات المتكررة
عينة واحدة أو عينتان أو عدة عينات	الارتباط بين القياسات أو العلاقة بين المتغيرات "دراسات ارتباطية"	مجموعة واحدة ذات اختبار قبلى أو بعدى أو عدة اختبارات	اسمية	معامل ارتباط فاى– معامل التوافق – معامل الاقتران الرباعى
			رتبية	معامل ارتباط سبيرمان– معامل ارتباط كندال
			فترية	معامل ارتباط بيرسون – الارتباط القانونى – الارتباط المتعدد
عينة واحدة أو عينتان أو عدة عينات	"دراسات تنبؤية " للمتغيرات أو عضوية الجماعة	مجموعة واحدة أو عدة مجموعات مع عدة اختبارات	فترية	تحليل الانحدار بأنواعه المختلفة– السلاسل الزمنية
				التحليل التمييزى بأنواعه المختلفة
عينة واحدة أو عينتان أو عدة عينات	" دراسات عاملية" البناء العاملى	مجموعة واحدة أو عدة مجموعة مع عدة اختبارات	فترية	التحليل العاملى الاستكشافى – التحليل العاملى التوكيدى