**جامعة: دمشق كلية الاقتصاد ماجستير: تسويق**

**تحليل البيانات باستخدام اساليب الاحصاء المعلمي**

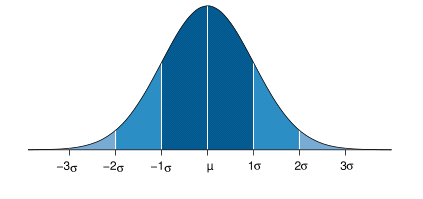
**إعداد الطلاب:**

**أيمن الشعباني**

**باسم شنو**

**عبد الغني قلعجي**

**باشراف الدكتور : سليمان عوض**



***2009-2010***

**مقدمة :**

تأتي مرحلة تحليل البيانات بعد عملية جمع البيانات وتبويبها ويقصد بالتبويب هو تنظيم البيانات في شكل مرتب يسهل من عملية تحليلها ومن أبرز أمثلة التبويب هو وضع الإجابات داخل جداول لتسهيل عرضها و تحليلها .

ويحتاج الباحثون إلى تحليل البيانات التي تم التوصل إليها لأسباب متعددة من هذه الأسباب أن اختيار اسلوب التحليل المناسب يزيد من قدرة الباحث على تفسير المتغيرات المؤثرة في الظاهرة لا سيما و ان كان عددها كبيرا . فمثلاً يؤدي استخدام أساليب التحليل الوصفية و خاصة التحليل العاملي إلى تمكين الباحث من التحليل المنطقي لتاثير المتغيرات المتعددة على ظاهرة البحث .

كذلك يسمح تحليل البيانات بالوقوف على مدى جوهرية تاثير بعض المتغيرات على ظاهرة فمثلا قد يظهر من الأرقام أن المتغير X يؤثر في الظاهرة Y.ويحتاج الباحث إلى معرفة مدى جوهرية هذا التأثير وهو ما تسمح به أساليب التحليل .وفضلا عن ذلك فهناك العديد من اساليب تحليل البيانات التي تحدد تأثير تفاعل أكثر من متغير على ظاهرة البحث . تأتي أهمية تقدير تاثير التفاعل عند اتخاذ القرارات التسويقية المناسبة بشأن عناصر المزيج التسويقي أو غيرها من العوامل المؤثرة في الخطة التسويقية للشركة.

واخيراً فإن أساليب تحليل البيانات تمكن الباحثين من تقدير بيانات المجتمع من واقع بيانات العينات الاحتمالية المأخوذة منه.

وفيما يلي سوف نسعى إلى تقديم فكرة مبسطة عن أساليب تحليل البيانات الأكثر شيوعاَ.

**أولا-محددات اختيار اسلوب التحليل:**

يتوقف اختيار اسلوب تحليل البيانات المناسب على مجموعة من الاعتبارات هي:[[1]](#footnote-1)

**1-نوع البيانات المتاحة:**

وتنقسم البيانات المتاحة لغرض تحليلها لأربع انواع :

أول نوع هو البيانات الاسمية .وهي تعني تقسيم الظاهرة إلى مجموعات لكل منها اسمها .وتتميز المفردات داخل كل مجموعة بالتشابه أو التساوي.

ثاني نوع هو البيانات الترتيبية .وهو يعني تقسيم الظاهرة إلى مجموعات بحيث تتشابه المفردات داخل كل مجموعة فضلاً عن امكانية ترتيب هذه المجموعات .

ثالث نوع هو البيانات القياسية .ويعني تقسيم الظاهرة إلى مجموعات تتميز المفردات داخل كل منها بالتشابه ,ويمكن ترتيب هذه المجموعات وتتساوى المسافات فيما بينها.

رابع نوع من البيانات هو النسب وفي هذه الحالة يمكن تقسيم المفردات إلى مجموعات لكل منهم اسم ويمكن ترتيب هذه المجموعات و اجراء التحويل الخطي للمسافات فيما بينهما كما يمكن تحديد نسبة البيان للآخر نظراً لوجود صفر طبيعي له.

وتختلف أساليب التحليل المستخدمة وفقا لنوعية البيانات المتاحة للتحليل وعموما يناسب البيانات القياسية والنسبية مستوى أرفع من أساليب التحليل التي تستخدم في حالة البيانات الاسمية والترتيبية .ولكن إذا كانت البيانات المتاحة للتحليل اسمية ونسبية مثلا فيضطر الباحث في هذه الحالة إلى معاملتها على أنها بينات اسمية حتى ينطبق عليها اسلوب التحليل .وهكذا يجوز معاملتة البيانات القياسية والنسبية على انها بينات اسمية ولكن لا يجوز العكس.

**2- تصميم البحث:**

يراعى عند تصميم البحث عدة اعتبارات .اول هذه الاعتبارات مدى استقلال العينة.وذلك لأن تحليل البيانات الناتجة عن عينة واحدة يختلف عن تحليل البيانات الناتجة عن عينات مستقلة .وثاني هذه الاعتبارات هو عدد المجموعات الناتجة.فمثلاَ يمكن استخدام التحليل الاحصائي أو اختبار Z أو اختبار "T" إذا كان عدد المجموعات اثنين .ولكن إذا زاد عدد المجموعات عن اثنين فيمكن استخدام اساليب تحليل اخرى مثل تحليل التباين ,وثالث اعتبارات تصميم البحث هو عدد المتغيرات المراد قياس تأثيرها على الظاهرة .

ويختلف أسلوب التحليل بإختلاف عدد المتغيرات .فمثلاً إذا كان عدد المتغيرات كبيرا مثل عشرة متغيرات فيمكن في هذه الحالة استخدام التحليل العاملي.

و آخر هذه الاعتبارات هو مدى القدرة على التحكم في المتغيرات المؤثرة في الظاهرة .بمعنى ان ظاهرة البحث قد تتأثر بمتغير او أكثر بخلاف المتغيرات التي ييتم قياس تاثيرها . وقد يصعب في بعض الحالات قياس تأثير جميع المتغيرات المؤثرة في الظاهرة.

**3-الشروط الخاصة بإستخدام أسلوب التحليل:**

يجب توافر عدة شروط لاستخدام كل اسلوب من أساليب تحليل البيانات .فمثلا يجب أن تأخذ النتائج شكل أرقام وليست نسب مؤية حتى يمكن أسلوب كذلك إذا كان حجم العينة صغيرا مثلاً فيمكن استخدام اختبار "T" وهكذا.

**ثانيا-الأهداف الرئيسية لتحليل البيانات:**

هناك ثلاثة أهداف لتحليل البيانات وهي: التعرف على البيانات ,واختيار مدى جودة البيانات,ثم اختبار فروض البحث التي تمت تنميتها في المراحل الأولى للبحث.اما التعرف على البيانات ,فإنه سيعطي الباحث فكرة مبدئية عن جودة المقياس المستخدم ,وعن جودة عمليات الترميز و إدخال البيانات للحاسب ...إلخ.فابفتراض ان الباحث يستخدم مقياس مورجان من 1 إلى 7 لقياس احد العناصر ,ولكنه وجد أن هذا العنصر قد رمز له ب 8 و ادخل إلى الكمبيوتر كذلك,فإن ذلك سوف يشير إلى وجود خطا في الترميز و إدخال البيانات ,كما أن الرقم 7 سوف يشير إلى اعلى قيمة متوقعة أثناء إجراء الاحصاءات الوصفية مثل التوزيع التكراري ...إلخ.أما الهدف الثاني (وهو اختبار مدى جودة البيانات) فمن الممكن تحقيقه عن طريق إجراء التحليل العاملي للبيانات ,والحصول على معامل كرنباخ الخاص بدقة البيانات ,إمكان الاعتماد على المقياس ...إلخ.أما الهدف الثالث وهو اختبار صحة فروض البحث ,فمن الممكن تحقيقه عن طريق اختيار مجموعة البرامج الاحصائية المناسبة لاختبار فروض البحث عن طريق اختيار مجموعة البرامج الاحصائية المناسبة لاختبار فروض البحث عن طريق الاختيارات الاحصائية المناسبة .وستشير نتائج التحليل الاحصائي إلى ما إذا كانت فروض البحث صحيحة أم لا.وسنناقش الآن كيفية استخدام تحليل البيانات لتحقيق كل هدف من هذه الأهداف الثلاثة.[[2]](#footnote-2)

**1-التعرف المبدئي على البيانات:**

يمكننا استخدام مقاييس النزعة المركزية والتشتت للتعرف المبدئي على البيانات .ذلك أنه من الممكن للباحث ان يستخدم المتوسط الحسابي والمدى و الانحراف المعياري والتشتت ليحصل على فكرة جيدة عن كيفية إجابة مفردات العينة على أسئلة الاستقصاء ومدى جودة المفردات المكونة للاستقصاء والمقياس بصفة عامة.فإذا لم تكن الاجابات عن سؤال معين منتشرة جيداً-مدى انتشار ضيق مثلا-فإن هذا يشير إلى اختلافات محدودة بين الاجابات عن ذلك السؤال ,مما قد يعني ان صياغة السؤال لم تكن جيدة ,وإلى ان أعضاء العينة لم يفهموا المقصود من السؤال.ومن الممكن التعرف على وجود التحيز في الإجابة في هذه المرحلة إذا كانت إجابات مفردات العينة على جميع الأسئلة متشابهة-بمعنى اختيار نقطة معينة على المقياس بالنسبة لجميع الأسئلة-4أو 3 مثلاً.ومن الممكن الحصول على اعلى و أدنى قيمة في اجابات افراد العينة ,وعلى المدى ,والمتوسط الحسابي والانحراف المعياري ,والتشتت وغير ذلك من الاحصاءات التي تشير إلى مدى توزع الإجابات على المقياس وما إذا كان ذلك التوزع مرضيا أم لا.تذكر أنه إذا لم يكن هناك اختلاف بين الاجابات فلن يوجد تشتت يشرحه الباحث ! وغالبا ما يفوم الباحث بالحصول على قائمة طويلة لمقاييس النزعة المركزية –المدى والتشتت وغيرهما من الاحصاءات لكل سؤال تقيس متغيرات البحث المستقلة والتابعة,خاصة إذا كان مقياس البحث جديداً لم يسبق استخدامه في عدد كاف من البحوث للتأكد من صلاحيته ,.ودقته ,وإمكان الاعتماد عليه.

وبالإضافة إلى الحصول على التوزيعات التكرارية والانحرافات معيارية وغيرها ,فقد يكون من المهم ان نتعرف على مدى ارتباط متغيرات البحث ببعضها,وذلك بالحصول على قائمة معاملات الارتباط بين هذه المتغيرات.

ويزيد تأكد الباحث من جودة البيانات الثقة في كل خطوات التحليل التالية وفي النتائج التي نحصل عليها من التحليل .ولذلك فإن التعرف على البيانات يعتبر ضرورة بالنسبة لجميع أنواع التحليل .وبناء على تلك المعرفة المبدئية يحدد الباحث مدة حاجته إلى مزيد من التحاليل الاحصائية الأخرى للتعرف على جودة البيانات.

**اختبار جودة البيانات:-2**

يمكننا الآن قياس دقة المقياس وصلاحيته ,لنتعرف على جودة البيانات التي استخدناها للحصول عليها.

**الدقة\امكان الاعتماد على المقياس:**

من الممكن التعرف على دقة المقياس عن طريق قياس كل من التماسك والاستقرار الموجودين بين مكونات المقياس .أما التماسك فإنه يشير إلى مدى الترابط بين المفردات التي تقيس مفهوما معينا ,بحيث تكون مع بعضها مجموعة واحدة .ويستخدم معامل كرنباخ للتعرف على قوة الارتباط الايجابي بين مفردات الاستبيان,وهو بالتالي مقياس لدقة الاستبيان.ويعتبر معامل كرنباخ مقياسا لقوة الارتباط الداخلي بين المتغيرات التي تقيس مفهوما معينا.زكلما اقترب هذا المعامل من 1 كلما كان ذلك دليلا على قوة التماسك الداخلي للمقياس.وهناك مقياس آخر لتماسك المقياس هو المقياس الذي يعرف بمعامل الثقة في نصفي المقياس .

من الممكن أيضا التأكد من ثبات نتائج المقياس باستخدام نموذجين متوازيين من المقياس أو عن طريق استخدام المقياس أكثر من مرة .فإذا كانت درجة الارتباط عالية بين نموذجي المقياس ,فإننا نكون قد تاكدنا من ثبات المقياس .كما يمكن التأكد من ثبات النتائج أيضا إذا كانت درجة الارتباط عالية بين نتائج القياسين الأول و الثاني-أي اللذين تم أخذها في وقتين مختلفين.

**الصلاحية**:

يمكن التأكد من توفر الصلاحية العاملية عن طريق اجراء التحليل العاملي على البيانات .وسوف تؤكد نتائج التحليل ما إذا كانت أبعاد النظرية ظاهرة بوضوح ام لا.كما يمكن التاكد من توفر صلاحية التقارب عندما يكون هناك ارتباط عال بين اجابات مصدرين مختلفين للبيانات ,مثل تشابه اجابات المشرفين ومرؤوسيهم عند قياس مدى معرفتهم لنظم التعويض عن المعمل الموجودة بالشركة ,ومن الممكن التأكد من صلاحية التمايز عندما يثبت عدم وجود ارتباط بين مفهومين من المفترض عدم وجود ارتباط بينهما مثل الشجاعة والأمانة ,والقيادة والحافز,والاتجاه والسلوك.

**اختبار صحة فروض البحث:-3**

بعد تجهيز البيانات للتحليل بالتخلص من الاجابات الخاطئة ,والتأكد من جودة البيانات فإن الباحث يكون مستعدا لاختبار مدى صحة فروض البحث .

**ثالثا- الأساليب الاحصائية لتحليل البيانات :**

هناك نوعان من الاساليب الاحصائية المتبعة وهما :**الاساليب الاحصائية الوصفية** التي تصف الظاهرة أوالمشكلة موضوع البحث ولا تتطرق إلى التعمق بدراسة الظاهرة و التعرف على بعض نواحي العلاقة او الارتباط بمتغيرات الدراسة .ومن أهم الأساليب الاحصائية الوصفية التكرار و الوسط الحسابي و المنوال والوسيط والانحراف المعياري ومقاييس التشتت.

**الاساليب الاحصائية الاستدلالية** فتسخدم في حال كون الدراسة ارتباطية وتهدف إلى التعق في دراسة الظاهرة ووضع فرضيات وفحصها بشكل إحصائي فإن الباحث يلجأ إلى استخدام مقاييس الارتباط مثل اختبار سبيرمان وبي رسون و الانحدار .كما قد يلجأ إلى بعض مقاييس الاختلاف مثل اختبار ** .[[3]](#footnote-3)**

وتقسم الاساليب الاحصائية الاستدلالية إلى أساليب احصائية معلمية ,وأساليب احصائية لا معلمية **.**

الاساليب اللامعلمية تستخدم في حالات لا يكون فيها نوع التوزيع الاحتمالي للمجتمع الاحصائي الذي سحبت منه عينة الدراسة معروفا ً أو في حالة عدم تحقق شرط كون التوزيع النظري للمجتمع الاحصائي طبيعياً . وهي تناسب البيانات الاحصائية الكمية و النوعية و المعطاة في صورة اسمية او رتبية .

أما الاساليب المعلمية فهي تتطلب افتراضات معينة حول المجتمع الاحصائي الذي سحبت منه عينة الدراسة منها:

1-أن يكون توزيع المجتمع طبيعيا ً .

2-أن يكون التباين متجانس.

وسنأتي على شرحها لاحقا ,ً سنركز في هذا البحث على الاساليب المعلمية.ومن ضمن الاساليب المعلمية الاساليب الوصفية في التحليل وسنبدأ عندها:

**الاساليب الاحصائية الوصفية :**

تتكون هذه المجموعة من ثلاث فئات من الأساليب الاحصائية هي : التكرار و الجداول المتقاطعة,وقياس النزعة المركزية ,والتشتت

# التكرار والجداول المتقاطعة:

**1-التكرار:[[4]](#footnote-4)**

يعرف التكرار بأنه عدد المرات التي تتكرر فيها الظاهرة أو مشاهدة معينة فمثلا إذا كان المتغير ,أنثى )فإن التكرار يعني في هذه الحالة عدد الذكور وعدد الاناث الخاص بمجتمع الدراسة أو العينة المختارة .ويستخدم التكرر بشكل أساسي في حالة المقاييس الاسمية مثل الأسئلة المتعلقة بالجنس والدين والحالة الاجتماعية.....إلخ.

كما يمكن استخدامه في حالة المقاييس التفاضلية مثل الأسئلة التي يطلب فيها إلى المبحوث إعطاء رأيه حول موضوع معين ويطلب إليه الإجابة عن ذلك ضمن خيارات تتدرج من موافق بشدة إلى غير موافق بشدة أو مهم جدا ً إلى غير مهم على الإطلاق.

ويستخدم في العادة مع التكرار النسبة المؤية لكل فئة التي تبين كل فئة إلى المجموع الكلي ,كما يستخدم أحيانا ً النسبة المؤية التراكمية التي تبين نسبة مجموع الفئات السابقة إلى المجموع الكلي .ولتوضيح التكرار ومجالات استخدامه نورد الأسئلة التالية المتعلقة بدراسة أجريت على عينة تتكون من 25 موظف إداري في مؤسسات الدولة المختلفة:

-المستوى التعليمي:

وعند فرز الإجابات كان المستوى التعليمي للعينة كما يلي:

|  |  |  |  |
| --- | --- | --- | --- |
| النسبة المؤية التراكمية | النسبة المؤية | العدد(التكرار) | المستوى التعليمي |
| 4% | 4% | 1 | ثانوية عامة أو اقل |
| 16% | 12% | 3 | دبلوم كليات مجتمع |
| 48% | 32% | 8 | بكالوريوس |
| 64% | 16% | 4 | دبلوم عالي بعد البكالوريوس |
| 88% | 24% | 6 | ماجستير |
| 100% | 12% | 3 | دكتوراه |
|  | 100% | 25 | المجموع |

نلاحظ من الجدول أعلاه أن العامود الثاني يظهر التكرار أو عدد أفراد العينة عند كل مستوى تعليمي والعامود الثالث يظهر نسبة عدد أفراد العينة عن كل مستوى تعليمي والعامود الأخير يعطي النسبة التراكمية ,فمثلا ً نلاحظ أن عدد الذين يحملون درجة البكالوريوس يبلغ 8 أشخاص وأنهم يشكلون ما سنبته 32% من العينة المختارة وأن ما نسبته 48%من العينة تحمل درجة بكالويوس أو أقل.

ولتوضيح كيفية استخدام التكرار في حالة الأسئلة التفاضلية نفترض ورود السؤال التالي مع الخيارات التالية ضمن الاستنبانة الموجهة للمجموعة السابقة نفسها من الموظفين:

أهمية الدورات التدريبية في رفع مستوى الأداء الوظيفي للعاملين في المؤسسات الحكومية ؟

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 5 | 4 | 3 | 2 | 1 |
| مهم جدا ً | مهم | متوسط الأهمية | غير مهم | غير مهم على الاطلاق |

وعند فرز إجابات المبحوثين ظهرت النتائج التالية :

|  |  |  |  |
| --- | --- | --- | --- |
| النسبة المؤية التراكمية | النسبة المؤية | العدد(التكرار) | الفئة |
| 32% | 32% | 8 | **مهم جدا ً** |
| 56% | 24% | 6 | **مهم** |
| 68% | 12% | 3 | **متوسط الأهمية** |
| 88% | 20% | 5 | **مهم غير** |
| 100% | 12% | 3 | **غير مهم على الإطلاق** |
|  | 100% | 25 | **المجموع** |

يمكن من خلال الجدول أعلاه استنتاج بعض الحقائق منها مثلا أن أفراد العينة الذين يعتقدون بأن الترتيب مهم في رفع الكفاءات يبلغ 6 أشخاص وهم يشكلون ما نسبته 24% من حجم العينة الكلي وأن نسبة الذين يعتقدون بأن هذا العامل مهم او مهم جدا ً (النسبة المؤية التراكمية )يبلغ 56% من المجموع الكلي.

**2-الجداول المتقاطعة:**

تعتمد الجداول المتقاطعة على مبدأ التكرار الا أنها تعطي معلومات أكثر عمقا ودلالة بالمقارنة مع ما تعطيه جداول التكرار , حيث تدمج الجداول المتقاطعة متغيرين أو أكثر وتبين مدى التقاطع في المعلومات الواردة في هذه المتغيرات ,فمثلا إذا كان احد الأسئلة يتعلق بالجنس(ذكر او انثى) وكان لدينا سؤال يتعلق بالمستوى التعليمي فإن نتائج كل سؤال تعطي معلومة منفصلة ولا تظهر لنا مثلا كم عدد الاناث اللواتي يحملن درجة البكالوريوس وهل هناك اختلاف واضح في المستوى التعليمي بين الذكر والاناث .أما عند استخدام الجداول المتقاطعة لهذين السؤالين فإن القارئ يستطيع معرفة عدد كل من الاناث و الذكور عند كل مستوى تعليمي.

فاعتمادا على المعلومات الواردة في المثال السابق والخاص بالعينة المكونة من 25 موظف أظهرت النتائج أن المستوى التعليمي للعينة كان بالشكل التالي:

|  |  |  |  |
| --- | --- | --- | --- |
| النسبة المؤية التراكمية | النسبة المؤية | العدد(التكرار) | المستوى التعليمي |
| 4% | 4% | 1 | ثانوية عامة أو اقل |
| 16% | 12% | 3 | دبلوم كليات مجتمع |
| 48% | 32% | 8 | بكالوريوس |
| 64% | 16% | 4 | دبلوم عالي بعد البكالوريوس |
| 88% | 24% | 6 | ماجستير |
| 100% | 12% | 3 | دكتوراه |
|  | 100% | 25 | المجموع |

على افتراض أن توزيع العينة بحسب الجنس كان بالشكل التالي:

|  |  |  |
| --- | --- | --- |
| النسبة المؤية | العدد(التكرار) | الجنس |
| 52% | 13 | ذكر |
| 48% | 12 | انثى |
| 100% | 25 | المجموع |

وعند دمج متغيري الجنس والمستوى التعليمي ظهر الجدول المتقاطع للنتائج بالشكل التالي:

|  |  |  |  |
| --- | --- | --- | --- |
| المجموع | انثى | ذكر | المستوى التعليمي \الجنس |
| 1(4%) | 1(4%) |  | ثانوية عامة أو اقل |
| 3(12%) | 3(12%) |  | دبلوم كليات مجتمع |
| 8(16%) | 4(16%) | 4(16%) | بكالوريوس |
| 4(16%) | 2(8%) | 2(8%) | دبلوم عالي بعد البكالوريوس |
| 6(8%) | 2(8%) | 4(8%) | ماجستير |
| 3(12%) |  | 3(12%) | دكتوراه |
| 25(100%) | 12(48%) | 13(52%) | المجموع |

نلاحظ من الجدول ان المعلومات التي تم الحصول عليها عند دمج المتغيرين الجنس والمستوى التعليمي قد أصبحت أكثر تعبيرا ودلالة مما هو الحال عليه في حالة عرض كل متغير وحده فمثلا نلاحظ أنه لا يوجد ضمن حملة الدكتوراه أيه أنثى كما أن نسبة الذكور الذين يحملون الماجستير تشكل 16% من العينة في حين تبلغ نسبة الاناث من حملة الماجستير 8% .

# 2-مقاييس النزعة المركزية:

إن معظم القيم لمختلف الظواهر الطبيعية تتركز عادة في الوسط أو قريبة منه , ومقاييس التمركز أو التوسط لأي مجموعة من البيانات التابعة الظاهرة هي تلك المقاييس التي تبحث في تقدير قيمة تتمركز حولها أغلبية البيانات بحيث تمثلها أفضل تمثيل وهناك ثلاث مقاييس احصائية مهمة وشائعة الاستخدام ,ويمكن أن تستخدم لتمثيل البيانات الاحصائية وحسب نوعية البيانات ,فيمكن أن يكون الوسط الحسابي هو من اهم المقاييس الاحصائية لكونه يستخدم جميع البيانات الاحصائية ,أما المقياس الثاني فهو الوسيط ويتعتبر من المقاييس المهمة و المستخدم بشكل واسع جداً وخاصة عندما يكون قسم من البيانات كبيرة جدا أو صغيرة جدا, او ما تسمى في الاحصاء بالقيم الشاذة ,أما المقياس الثالث فهو المنوال والذي يستخدم بشكل واسع ويعتبر أيضا من المقاييس المهمة وخاصة عند البيانات الوصفية و تستخدم التكرارات الاحصائية .كما وسيتم في هذا الفصل التعرف على المدى, الانحراف المتوسط ,التباين و الانحراف المعياري كمقاييس للتشتت والتي تقيس التغيرات الموجودة في البيانات والتي بدورها مع مقاييس النزعة المركزية تعطي صورة واضحة عن البيانات .وسيتم عرض جميع هذه المقاييس كالآتي:[[5]](#footnote-5)

**الوسط الحسابي:-1**

الوسط الحسابي هو أحد أهم مقاييس النزعة المركزية و أكثرها استخداما وشيوعا و عادة ما يسمى بالمعدل ويفضل على جميع مقاييس النزعة المركزية لكونه يستعمل جميع البيانات ويستخدم الصيغ الرياضية .[[6]](#footnote-6)

**الوسط الحسابي للبيانات الخام أو ما يسمى بالقيم غير المبوبة:-1-1**

الوسط الحسابي للبيانات الخام هو مجموع القيم مقسوما على عددها . ولهذا ففي حالة وجود القيم المتطرفة الكبيرة جدا أو الصغيرة جدا (القيم الشاذة) فيفضل استخدام الوسيط .ويمكن أن يحسب الوسيط كما يلي:

الوسط الحسابي لمجموعة "N" من المشاهدات او القيم :

X هو X1,X2……………….Xn

ويمكن ايجاده :

=

للقيم من :

I=1,2, ……,n

مثال(1):أوجد الوسط الحسابي للبيانات التالية:

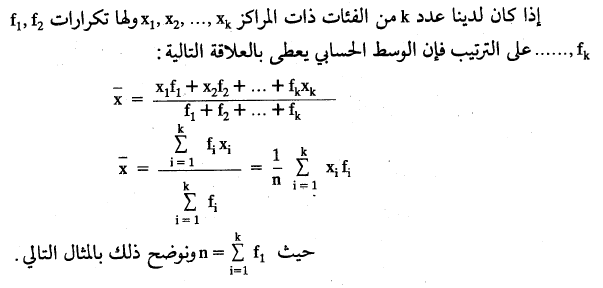
75,74,72,71,70,69,68,67,66,63

الحل:

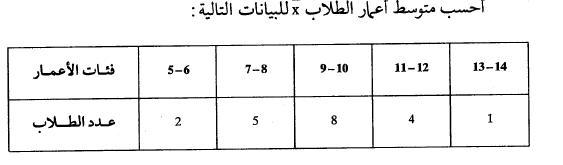
X==

**الوسط الحسابي للبيانات المجمعة أو ما يسمى بالقيم المبوبة:-2-1**

عندما تكون البيانات موضوعة على شكل جدول تكراري سوف يتم استخدام مركز الفئة iX والذي يساوي مجموع الحدين الأدنى و الأعلى مقسوما على 2 لتمثيل تلك الفئة.وعندما تتوزع البيانات توزيعا طبيعيا ً فإن جميع المقاييس الوسط الحسابي و الوسيط و المنوال لها نفس القيمة .أما عند رسم المنحنى التكراري ويكون الطرف الأيسر أطول من الطرف الأيمن ,أي أطول نحو القيم الصغيرة فإن قيمة الوسيط تكون أكبر من قيمة الوسط الحسابي أما إذا كان الطرف الأيمن أطول من الطرف الأيسر أي أطول نحو القيم الكبيرة فإن قيمة الوسط الحسابي تكون أكبر من الوسيط.



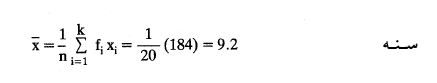
**مثال:**



الحل:ولسهولة الحل نضع الجدول التالي:

|  |  |  |  |
| --- | --- | --- | --- |
| الفئات | X مراكز الفئات | fالتكرار | fx |
| 5-6 | 5.5 | 2 | 11 |
| 7-8 | 7.5 | 5 | 37.5 |
| 9-10 | 9.5 | 8 | 76.0 |
| 11-12 | 11.5 | 4 | 46.0 |
| 13-14 | 13.5 | 1 | 13.5 |
| المجموع |  | 20 | 184 |

وبالتعويض بالقانون السابق نحصل على:



**بعض خصائص الوسط الحسابي:[[7]](#footnote-7)**

من اهم خصائص الوسط الحسابي مايلي:

**1-مجموع انحرافات القيم عن وسطها الحسابي تساوي الصفر,أي أن :**

أولا ً –في حالة البيانات غير المبوبة :

ثانيا ً في حالة البيانات المبوبة:

**2-مجموع مربعات انحرافات القيم عن وسطها الحسابي أقل ما يمكن** ,اي أقل من مجموع مربعات انحرافات القيم عن أية قيمة غير الوسط الحسابي ,أي ان :

حيث B أي قيمة أخرى غير X.

**3- عند اضافة كمية ثابتة أو طرح كمية ثابتة ولتكن C إلى أو من كل قيم من القيم فسوف يكون الوسط الحسابي الجديد هو الوسط الحسابي القديم مضافا إليه أو مطروحا منه الكمية الثابتة C** ,أي فذلك يعني أن: و أن القيم الجديد هي ان اذا افترضنا أن القيم القديمة هي

di=Xi+c

**عند ضرب كل قيمة من العينة في كمية ثابتة"C" فإن:-4**

الوسط الحسابي للقيم الجديدة هو الوسط الحسابي للعينة أي القيم القديمة مضروبة في الكمية الثابتة C , أي بافتراض أن  هي قيم العينة (القيم القديمة), و ان  هي القيم الجديدة والتي نتجت عن ضرب القيمبالقيمة الثابتة C فإن:

Yi=cxi

وبالتالي فإن:

Y=CX

**5-الوسط الحسابي لجموع عدة عينات هو مجموع الأوساط الحسابية لهذه العينات:**

فلو فرضنا أن لدينا ثلاث عينات X,Y,Z.

فإن مجموع العينات هو :

Ci=iY+Zi+Xi

وإن الوسط الحسابي لهذه العينات هو :

C=Y+Z+X

**بعض مميزات الوسط الحسابي:[[8]](#footnote-8)**

1-مقياس سهل حسابه ويخضع للعمليات الجبرية بسهولة

2-يأخذ في الاعتبار جميع القيم نحل الدراسة.

3-أكثر المقاييس فهما في الاحصاء.

**بعض عيوب الوسط الحسابي:**

1-يتأثر بالقيم المتطرفة(وهي القيم الكبيرة جدا ً أو الصغيرة جدا ً مقارنة ببقية القيم)

2-يصعب حسابه في حالة الجداول التكرارية المفتوحة حيث يتطلب معرفة مركز كل فئة.

3-يصعب حسابه في حالة البيانات الوصفية.

**الوسط الحسابي الموزون:[[9]](#footnote-9)**

إذا كان لكل قيمة من المشاهدات Xi وزن خاص بها يتناسب مع أهميتها ويرمز له iW فيكون الوسط الحسابي في هذه الحالة بالشكل:

X=

**الوسيط :-2**

هو المقياس الثاني من مقاييس النزعة المركزية في الأهمية وبحسب إذا تم ترتيب البيانات حسب حجمها تصاعديا أو تنازليا .و الوسيط يعرف بأنه مقياس الموقع لأن قيمته تعتمد على موقعه في البيانات وتوجد بعض مقاييس المواقع الأخرى مثل الربيع والمئين.وهو يكون القيمة التي تقع في وسط البيانات[[10]](#footnote-10) , وهذه الميزة من ناحية جيدة كونه لا يتأثر بالقيم الكبيرة أو الصغيرة أو ما تسمى بالقيم الشاذة وهذه الخاصية تميزه على الوسط الحسابي ومن ناحية أخرى فهذه الخاصية تعتبر الأخذ على الوسيط هو كونه يستخدم هذه القيمة فقط ويهمل جميع المعلومات في القيم الباقية وهنا يفضل عليه الوسط الحسابي إذا لم تكن هناك قيم شاذة .

**الوسيط للبيانات غير المبوبة :**

لحساب الوسيط بالنسبة للبياناتا لخام ,نقوم بترتيب البيانات من أصغر قيمة إلى أكبر قيمة فإذا كان عدد المشاهدات n فردي ,يكون الوسيط هو القيمة الوسطية ,أما إذا كان عدد المشاهدات n زوجي ,فيكون الوسيط القيمتين الوسطيتين.

2Q=

**الوسيط للبيانات المبوبة:**

عندما تكون اليبانات الموجودة في جدول توزيع تكراري فإن الفئة التي تقابل رتبة الوسيط والتي تساوي نصف مجموع التكرارات في جدول التوزيع التكراري المتجمع الصاعد ,تسمى فئة الوسيط أو الفئة الوسيطية ويمكن إيجاد قيمة الوسيط من خلال فئة الوسيط هذه باستعمال القانون التالي:

M=L+[]w

حيث أن : L: تمثل الحد الأدنى الحقيقي لفئة الوسيط

W: تمثل عرض فئة الوسيط.

: تمثل مجموع التكرارات الكلية.

تمثل تكرار فئة الوسيطf

F تمثل تكرار المتجمع الصاعد الذي يقابل الحد الأدنى لفئة الوسيط.

**المنوال:-3**

يمثل المنوال قيم المشاهدات أو الحوادث الاكثر تكرارا ً أو حدوثا ,وفي حالة كون المشاهدات معروضة على شكل فئات تكرارية فان المنوال يكون في هذه الحالة الوسط الحسابي للفئة الأ كثر تكرارا ً .ويستخدم المنوال في كثير من الدراسات الادارية للتعبير عن اكثر العناصر أو المشاهدات تكراراً,فقد يتم استخدامه للتعبير عن أكثر اصناف العطور مبيعاً مثلاً[[11]](#footnote-11).ولتوضيح كيفية استخراج المنوال سنورد المثال التالي الذي يمثل احد الأسئلة التي وردت ضمن احد الدراسات التي اجريت  
على عينة مكونة من 50 موظف يحملون درجة البكلوريوس ويعملون لدى الشركات الصناعية,وقد كان السؤال يتعلق بالتخصص ,وقد ظهرت النتائج كما يلي:[[12]](#footnote-12)

**المنوال حسب افراد عينة دراسة حسب التخصص**

|  |  |
| --- | --- |
| العدد(التكرار) | المستوى التعليمي |
| 5 | ادارة عامة |
| 10 | ادارة اعمال |
| 20 | محاسبة |
| 6 | اقتصاد |
| 7 | لغة انكليزية |
| 2 | كمبيوتر(حاسوب) |
| 50 | المجموع |

نلاحظ أن أكثر التخصصات تكراراً هو المحاسبة وبالتالي فإنه يكون هو المنوال لهذه الحالة.

وبالتالي يمكن ان نستخدم القانون التالي عندما تكون البيانات مبوبة:[[13]](#footnote-13)

**M0=L+[] w**

حيث أن :

L تمثل الحد الأدنى الحقيقي لفئة المنوال.

W تمثل عرض فئة المنوال.

1d الفرق بين تكرار الفئة المنوالية وتكرار الفئة السابقة لها.

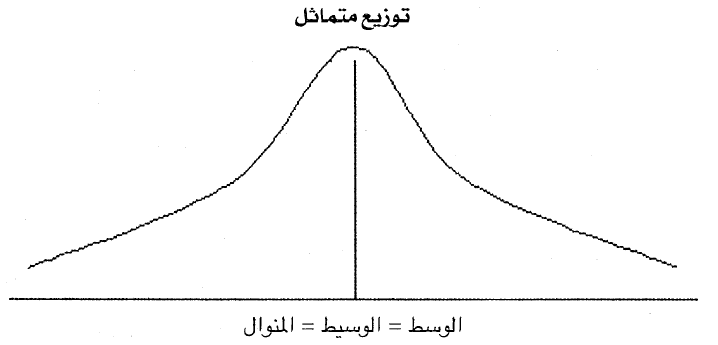
2d الفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها.

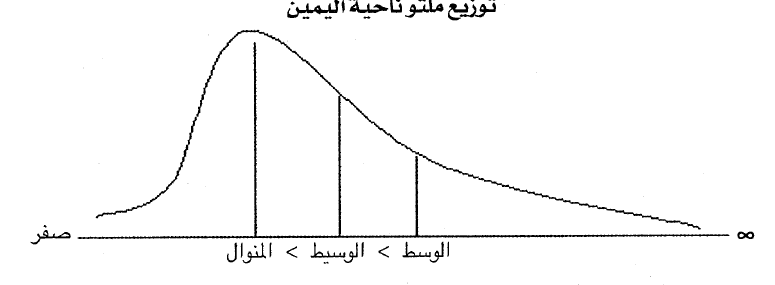
**العلاقة بين الوسط الحسابي والوسيط والمنوال:[[14]](#footnote-14)**

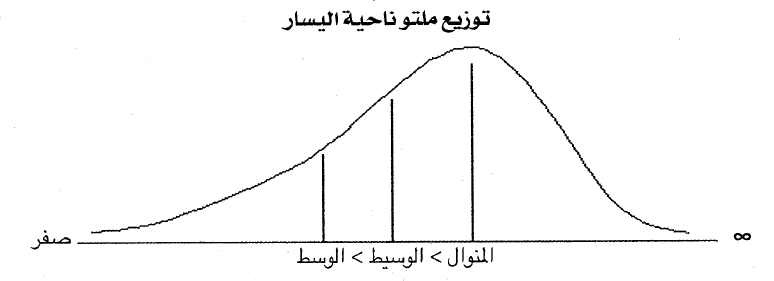
إذا كان لمجموعة من البيانات منوال واحد فإن المتوسط الحسابي والوسيط والمنوال تربطها إحدى العلاقات التالية:

1

أ-إذا كان التوزيع التكراري متماثلا ً فإن المنحنى سيكون متمائثلا وله ثمة واحدة وشكله يشبه شكل الجرس(شكل التوزيع الطبيعي) وفي هذه الحالة يكون المتوسط الحسابي مساويا للوسيط ويساوي المنوال ويساوي مركز الفئة التي يقابلها أكبر تكرار ,كما يتضح في الشكل التالي:

ب- إذا كان المتوسط الحسابي أكبر من الوسيط و المنوال ,فإن المنحنى التكراري سيكون ملتويا تاحية اليمين ,كما يتضح من الشكل التالي:

ج- إذا كان المتوسط الحسابي أقل من الوسيط والمنوال فإن المنحنى التكراري سيكون ملتويا ناحية اليسار كما يتضح من الشكل التالي:



ملحوظة :دائما تكون قيمة الوسيط ما بين قيمة المتوسط الحسابي وقيمة المنوال .وعليه إذا كان التوزيع التكراري وبالتالي المنحنى الذي يمثله ,ملتويا قليلا ُ فإن العلاقة التالية تكون صحيحة :

**الوسط الحسابي – المنوال =3(الوسط الحسابي –الوسيط)**

وقد وجد أن الوسيط تقع قيمته بين الوسط الحسابي والمنوال .

**الوسط الهندسي:[[15]](#footnote-15)-4**

الوسط الهندسي M.G لمجموعة من القيم nx .....,2x, 1x هو الجذر النوني لحاصل ضرب هذه القيم:



ويمتاز الوسط الهندسي عن الوسط الحسابي لأنه أقل تأثرا بالقيم الشاذة في البيانات لأنه معلوم رياضيا بأن الوسط الهندسي لمجموعة من القيم أقل من وسطها الحسابي,وعادة ما يحسب الوسط الهندسي باستخدام اللوغاريتمات كالتالي:



**الوسط الهندسي في حالة البيانات المبوبة:**

X1.x2………xk

في هذه الحالة يحسب الوسط الهندسي للفئات التي عددها ومراكزها هي

F1.f2……fk

والتي يقابلها بالترتيب تكرارات فإن الوسط الهندسي يعطى بالعلاقة التالية:

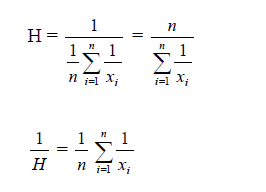


**5-الوسط التوافقي:[[16]](#footnote-16)**

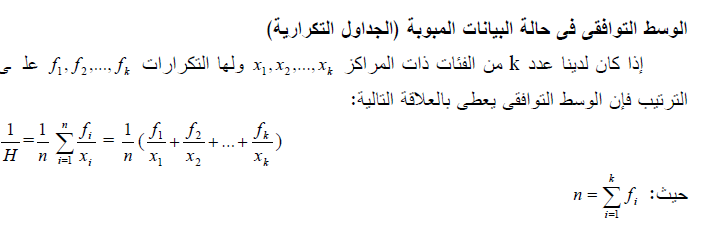
يستخدم الوسط التوافقي عندما يكون مقلوب المتغيرله دلالة كأن يعين نسبة بين متغيرين مرتبطين مثل السرعة بالنسبة للزمن .والوسط التوافقي لمجموعة من القيم هو مقلوب الوسط الحسابي لمقلوبات هذه القيم ,أي أن :

X1,x2,………..xn

H



من الناحية العملية يكون كالتالي :



**6-الربيعات والعشيرات والمئينات:[[17]](#footnote-17)**

مقاييس النزعة المركزية السابقة لها استخداماتها الخاصة بها وهذه المقاييس الربيعات والعشيرات و المئينات لها استخداماتها .فمثلا يستخدم الوسيط لتحديد القيمة التي تقع في منتصف البيانات او تقسم المساحة تحت المنحنى التكراري إلى قسمين متساويين.ويمكن ان تبرز الحاجة إلى تقسيمات أخرى للبيانات أو تقسيمات المساحة تحت المنحنى ولهذا فالربيعات تقسم البيانات إلى أربعة اقسام متساوية أو تقسم المساحة تحت المنحنى إلى أربعة أقسام متساوية وكذلك الحال بالنسبة للمئينات تقسم المساحة أو البيانات إلى مائة جزء متساوي وهذه التقسيمات لها استخدامات واسعة و كثيرة جدا فمثلا يمكن تقسيم مجموعة من الطلبة حسب المعدل إلى أربعة أقسام متساوية أو تقسيم المجموعة نفسها إلى مائة قسم ويمكن تحديد موقع كل طالب في هذه المجموعة حسب التقسيمات السابقة أو تقسيم مجموعة من الأشخاص حسب الدخل إلى أربعة أقسام أو مائة قسم متساوية وحسب التسلسل و يمكن تحديد أي ترتيب لأي من الأشخاص فالربيع الثاني 2Q هو نفسه العشير الخامس 5D و كذلك المئين الخمسون 50P.

ويمكن تعريف هذه التقسيمات كما يلي:

القيم الثلاث التي تقسم توزيع البيانات إلى أربع أجزاء متساوية تعرف على أنها الربيعات.

القيم التسعة والتسعون التي تقسم توزيع البيانات إلى مائة قسم متساوية تعرف بالمئينات.

القيم التسعة التي تقسم توزيع البيانات إلى عشرة أقسام متساوية تسمى بالعشيرات أو العشريات.

k

افترض أن هناك n من الأرقام مرتبة ترتيب تصاعدي .فيمكن إيجاد الربيعات والمئينات كما يلي:

موقع الربيع الأول 1Q تقابل ترتيب 4 / (1+n)

Q1=1/4(n+1)

موقع الربيع الثاني 2Q و هو الوسيط تقابل ترتيب بيانات4/ (1+n)2

Q2=1/2(n+1)

موقع الربيع الثالث 3Q تقابل ترتيب بيانات4/(1+n)3

4/3=3Q (n+1)

المدى الربيعي هو الفرق بين الربيع الثالث 3Q والربيع الأول 1Q:

=Q3-Q1

نصف المدى الربيعي هو الفرق بين الربع الثالث 3Q والربيع الأول 1Q:

=1/2(Q3-Q1).

المدى المئيني هو الفرق المئين التسعون والمئين العاشر.

P90-P10.

ويلاحظ أن الفائدة من تلك المئينات هو انها تعتمد بشكل كامل على النصف الأوسط من البيانات ولهذا لا تتأثر بالقيم المتطرفة الكبيرة والصغيرة.

# مقاييس التشتت:-3

لقد سبق لنا دراسة طرق عرض البيانات جدوليا وبيانيا والتعرف على أشكالها وتوزيعاتها المختلفة وكذلك دراسة مقاييس النزعة المركزية (المتوسطات)وذلك لوصف البيانات عدديا لهذه التوزيعات المختلفة .ولكن طرق عرض البيانات وحساب المتوسطات للمجموعات المختلفة من البيانات غير كافية للمقارنة بين هذه المجموعات .

حيث ان اقتصار وصف البيانات على استخدام مؤشرات النزعة المركزية لا يعطي صورة واضحة او كافية عن هذه البيانات .إذ من الممكن أن تجد عددا من التوزيعات التي لها نفس المتوسط او حتى تتساوي كافة مؤشرات النزعة المركزية وفي نفس الوقت تختلف كثيرا في درجة تشتيتها أو في أشكال توزيعها.[[18]](#footnote-18)

لذلك نشأت حاجة إلى ايجاد مقاييس تقيس درجة تجانس (تقارب)او تشتت (تباعد)مفردات البيانات بعضها عن بعض وهذه المقاييس هي ما تسمى بمقاييس التشتت,وهي كثيرة.

[[19]](#footnote-19)X,Y,Z **مثال**: لدراسة درجات ثلاث مجموعات مختلفة من الطلاب

X;59,61,62,58,60

Y;50,60,66,54,70

Z;19,65,46,78,72

وبحساب المتوسط الحسابي لثلاث مجموعات نجده يساوي 60 درجة لكل منها ولكن عند النظر إلى درجات المجموعة الأولى نجدها متقاربة ,ودرجات المجموعة الثانية أقل تقاربا من الأولى ,والمجموعة الثالثة أقل تقاربا من المجموعة الثانية.أي أن المجموعات المختلفة التجانس على الرغم من أن المتوسط الحسابي لثلاث المجموعات نفسه,وبالتالي فمقاييس النزعة المركزية غير كافية للمقارنة بين طبيعة البيانات الاحصائية.

وبعكس مقاييس النزعة المركزية ,التي تعد مقاييس نقاط او مستوى ,ولها وجود فعلي في التوزيع ,فإن مقاييس التشتت تعتبر مقاييس بعد او مسافة ,وليست قيما فعلية من بيانات التوزيع ,وبالتالي فهي لا يمكن أن تكون قيما سالبة أيضا ً ,فمثلا ً إذا كانت قيمة تشتت العمر لمجموعة من المبحوثين هي (3.5سنة) فإن هذا يعني أن أعمار هؤلاء المبحوثين تبتعد عن المتوسط (إما بالزيادة أو بالنقصان ) أو عن بعضها البعض ,كما أن القيمة (335 ) لايجب أن تكون بالضرورة ضمن قيم التوزيع.

ويمكن تقسيم مقاييس التشتت إلى مجموعتين هما :مقاييس تقيس تقارب أو تباعد القيم عن بعضها البعض وهي المدى و الانحراف المعياري (أو نصف المدى الربيعي).والأخرى مقاييس تقيس قرب أو بعد القيم من قيمة معينة كالمتوسط الحسابي مثلا ُ وهي:الانحراف المعياري.

**المدى:**

يعتبر المدى أسهل طريقة لقياس درجة التشتت للبيانات الكمية ,ويستخدم عندما يكون الهدف هو الحصول على مقياس سريع لمدى تشتت المفردات,دون الاهتمام الكبير بالدقة في القياس ,أو حينما يكونللمفردات المتطرفة أهمية خاصة مثل دراسات مراقبة جودة الانتاج ويعرف المدى على انه الفرق بين أعلى قيمة و أصغر قيمة ,ويعتمد بشكل كامل على القيمتين المتطرفتين .

المدى=أكبر قيمة – أصغر قيمة.

لكون المدى يعتمد على هاتين القيمتين فإنه يتأثر بهذه القيم بشكل كبير جداً وخاصة كون إحدى القيم أو الاثنين قيم شاذة وهذا الماخذ على المدى قلل من اهميته وكذلك من استخدامه مع سهولته الكبيرة.[[20]](#footnote-20)

أما في حالة البيانات المبوبة فإن المدى يعرف بأكثر من طريقة تذكر منها فيما يلي الطريقتي:

1-المدى=الفرق بين مركزي الفئة العليا والفئة الدنيا.

2-المدى =الحد الأعلى للفئة العليا مطروحا منه الحد الأدنى للفئة الدنيا .

**بعض خواص المدى:[[21]](#footnote-21)**

بسيط الحساب وسهل المفهوم,ولذلك فهو كثير الاستخدام في الأوساط العامة.-1

يستخدم في رسم الخرائط الاحصائية لمراقبة مطابقة الإنتاج للموصفات المطلوبة .-2

3-يمكن أن يحقق فائدة معقولة إذا تم استخدامه جنبا ً إلى جنب مع أحد مقاييس النزعة المركزية (المتوسط مثلاً)للمقارنة بين توزيعين أو أكثر من البيانات.

4-لايعتمد في حسابه على كل البيانات.

مقياس مضلل في حالة وجود قيم شاذة ومتطرفة.-5

لا يمكن حسابه من الجداول التكرارية المفتوحة.-6

هو طول أقصر فترة تحتوي كل البيانات.-7

لايقيس التشتت عن مقياس مركز معين.-8

ويمكن القول أيضا ً بأن كفاءة المدى تقل كثيرا ُ مع تغير طبيعة وحجم العينة .كما يعتبر مقياسا مضللا عندما يستخدم لمقارنة مجموعتين تختلفان في الحجم ,فزيادة حجم المجموعة ربما يزيد من احتمالية وجود قيم متطرفة ,إضافة إلى ان المدى ربما لا يفيد في إعطاء صورة عن شكل انتشاء البيانات ,خاصة عندما يتساوى مدى مجموعتين من البيانات على الرغم من ختلاف متوسط كل منهما.

**الانحراف الربيعي:[[22]](#footnote-22)**

للتغلب على بعض عيوب المدى ,والتي من أهمها تأثره بالقيم الشاذة وعدم إمكانية حسابه في حالة الجداول التكرارية المفتوحة وجد مقياس آخر لظاهرة التشتت وهو ما يسمى بالمدى الربيعي أو( نصف المدى الربيعي ) والفكرة التستسية في هذا المقياس هي ترتيب البيانات ثم حذف ربع القيم من المتطرفين حتى نتخلص من القيم الشاذة أو المتطرفة (أو الفئات المفتوحة) والاعتماد فقط على النصف الأوسط للقيم ,فيؤخذ المدى الواقع في هذا النصف الأوسط ,وفي مثل هذه الحالة تكون أكبر قيمة في البيانات هذا النصف هي الربيع الأعلى (الثالث)و أصغر قيمة الربيع الأدنى(الأول)والفرق بينهما يعطى ما يسمى بالمدى الربيعي,ويتم قسمة المدى الربيعي على 2 فنحصل على نصف المدى الربيعي (أو الانحراف الربيعي )وهو ما ستخدم كمقياس للتشتت.

**الانحراف المعياري:**

الانحراف المعياري هو ثالث أداة لقياس التشتت و اهم أدوات التشتت ,ويعرف بأنه الجذر التربيعي للتباين حيث تتبع لايجاده الخطوات الخاصة نفسها في ايجاد التباين ثم ايجاد الجذر التربيعي للتباين فيكون الناتج هو الانحراف المعياري.[[23]](#footnote-23)

يعرف التباين بأنع متوسط انحرافات القيم عن وسكها الحسابي ويرمز له بالرمز

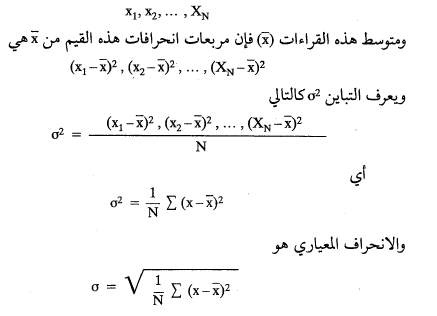
والجذر التربيعي للتباين ينتج ما يسمى بالانحراف المعياري وسوف نتطرق

لطرق حسابه في حالة البيانات المباشرة والبيانات المبوبة كما يلي:[[24]](#footnote-24)

**الانحراف المعياري للبيانات غير المبوبة:**

إذا كان لدينا قراءات من مجتمع احصائي عدد مفرداته هي:

N



σ

ويفضل عند حساب الانحراف المعياري أن يحسب التباين من المعادلة الأولى وبأخذ الجذر التربيعي للنتيجة النهائية نحصل على وفي حالة العينة التي حجمها المأخوذة من المجتمع فإن الانحراف المعياري في هذه الحالة يرمز له بالرمز والتباين ويعرف بقسمة مجموع مربعات الانحراف على بدلا من في حالة ويكتب كمايلي :

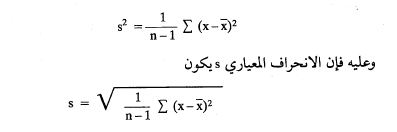
n

σ

s

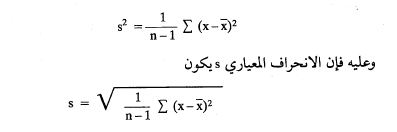
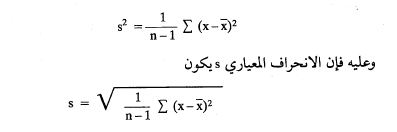
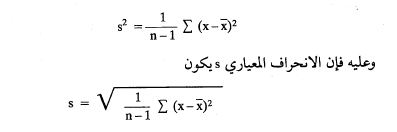
N

(n-1)



وذلك في حالة البيانات المباشرة غير المبوبة فإن يعطي تقديرا ً أحسن للانحراف المعياري للمجتمع الذي أخذت منه العينة .وإذا كان عدد مفردات العينة كبيرا ً (أكبر من 30 )فإن قيمة , متساويتان تقريبا من الناحية العملية.

σ



s

حيث ان أي مقياس للتشتت يجب أن لا يقل عن الصفر لذا لزم أخذ الجذر التربيعي الموجب للتباين لايجاد الانحراف المعياري.

**معامل الاختلاف :[[25]](#footnote-25)**

من المعلوم أن الانحراف المعياري لمجموعة من البيانات يأخذ وحدات البيانات نفسها فإذا كانت البيانات تمثل الأطوال مقيسة بالسنتميترات فإن الانحراف المعياري يكون بالسنتيمتر و لإذا كانت البيانات تمثل الأوزان فإنها تكون مقيسة بالكيلوغرام ويكون الانحراف المعياري مقيسا بالكيلوغرام فإذا أردنا مقارنة تجانس مجموعة من الأوزان أو تشتتها بمجموعة من الأطوال فلا يمكن استخدام الانحراف المعياري للمقارنة لأنه لا يمكن مقارنة السنتيمتر بالكيلوغرام لذا دعت الحاجة إلى ايجاد مقياس لا يعتمد على الوحدات وهذا المقياس هو ما يسمى بمعامل الاختلاف أو التشتت النسبي ويعرف كالتالي:

**المتغير المعياري والدرجات المعيارية (مقياس التمركز):**

نحتاج في بعض الأحيان إلى مقارنة مشاهدات مجاميع مختلفة , وهنا نحتاج لتحويل هذه المشاهدات إلى وحدات قياسية حتى نتمكن أن نقوم بالمقارنة بين المشاهدات وذلك باسنخدام الوسط الحسابي والانحراف المعياري لكل مجموعة من مجاميع المشاهدات وكمايلي:

باستخدام القانون التالي الخاص ب ( ) التي تمثل القيمة أو الدرجة المعيارية نجد أن :

Zi

حيث يمكن ايجاد قيمة Z لكل مجموعة من المجاميع باستخدام الوسط الحسابي الخاص بالمجموعة وكذلك الانحراف المعياري الخاص بالمجموعة أيضاً وهنا تصبح قيمة (Z ) خالية من الوحدات ولهذا نستطيع القيام بالمقارنة بين قيم Z التي كل واحدة تمثل مجموعة خاصة بها .[[26]](#footnote-26)

حيث أن : هي الدرجة المعيارية = Zi

هو الوسط الحسابي = Xi

الانحراف المعياري =iS

وعند أخذ مجموعة واحدة من البيانات تتكون من عدة مشاهدات وتم تحويل مشاهداتها إلى وحدات معيارية Zi وذلك بطرح الوسط الحسابي للمجموعة منها وتقسيمها على الانحراف المعياري للمجموعة فإن الوسط الحسابي لهذه المشاهدات المحولة إلى( iZ أي الوحدات المعياري) سوف يكون (1) و أن الانحراف المعياري لها سوف يكون (0) .أي أن:

σ

**مقاييس الالتواء:**

لقد سبق وتحدثنا عن طرق عرض البيانات جدولياً وبيانيا ً ثم مقاييس النزعة المركزية ومقاييس التشتت ولم نتعرض لأي مقياس يوضح درجة التواء المنحنى التكراري .ويقصد بكلمة التواء هو بعد المنحنى عن التماثل,ويكون منحنى التوزيع التكراري ملتويا نحو اليمين إذا كانت القيم متطرفة نحو اليمين تؤثر على الوسط الحسابي وتتجه نو اليمين ,وبذلك يكون الوسط الحسابي أكبر من الوسيط أما إذا كان التوزيع ملتويا نحو اليسار فإن الوسط الحسابي يكون أصغر من الوسيط.[[27]](#footnote-27)

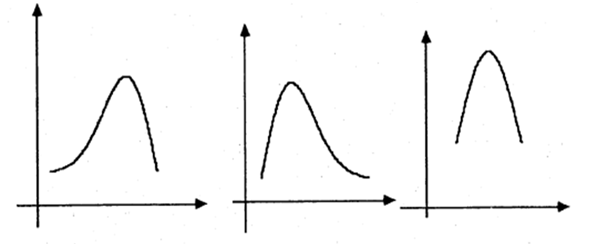
مقياس الالتواء له أشكال لمختلفة نذكر منها وهي :

كما تستخدم طريقة العزوم وتعطى من خلال العلاقة:

حيث:

وذلك في حالة البيانات المباشرة

وذلك في حالة البيانات المبوبة.



**ملتو نحو اليسار سالب الالتواء**

**ملتو نحو اليمين موجب الالتواء**

**متماثل الالتواء**

**القسم الثاني:**

**الاستدلال الإحصائي للعينات كبيرة الحجم**

**الفرق بين التوزيع الطبيعي والتوزيع الطبيعي المعياري [[28]](#footnote-28)**

خواص التوزيع الطبيعي :

المساحة التي يحصرها تساوي الواحد الصحيح

إن منحني التابع هذا متناظر حول النقطة وهي النقطة التي يتمركز عندها التوزيع أو يبلغ نهايته العظمى

في كل مرة نحتاج فيها إلى حساب المساحة أو الاحتمال لا بد من إجراء هذه التكاملات الشاقة والسبب يعود إلى اختلاف الأوساط الحسابية للمنحنيات وانحرافاتها المعيارية

ومن هنا جاءت فكرة البحث عن منحني طبيعي له شكل واحد

**التقدير الإحصائي وأنواعه :[[29]](#footnote-29)**

يعد تقدير مجتمع الإحصائي بدلالة إحدى العينات العشوائية المأخوذة منه من أحد الأسباب الرئيسية لتحليل العينات إذ يمكن التمييز بين نوعين من التقدير هما :

**التقدير النقطي** : هو عبارة عن قيمة واحدة فقط يأخذها الثابت الإحصائي المقدر بدلالة التابع الإحصائي المقابل والمحسوب من تلك العينة المسحوبة من المجتمع الإحصائي المراد تقدير أحد ثوابته الإحصائي أي تقدير الثابت بقيمة واحدة مثلا : يكون أفضل تقدير للوسط الحسابي للمجتمع الإحصائي هو الوسط الحسابي المحسوب من عينة عشوائية سحبت من ذلك المجتمع على اعتبار أن هناك احتمالا كبيرا ليكون الوسط الحسابي للعينة قريبا جدا من الوسط الحسابي للمجتمع غير المعلوم ،وما ينطبق على الوسط الحسابي ينطبق على الثوابت الإحصائي الأخرى

**صفات هذا التقدير**

**عدم التحيز:** يقال عن التابع الإحصائي بأنه غير متحيز للثابت الإحصائي إذا كان الثابت الإحصائي هو الأمل الرياضي للتابع الإحصائي

**فعالية التقدير :** فكلما قل التباين زادت فعالية التقدير ،أي التباين عن الثابت الإحصائي أقل ما يمكن

**التقارب:** ويقال عنتقدير أنه متقارب إذا كانت قيمة التابع الإحصائي تقارب قيمة الثابت الإحصائي كلما زاد حجم العينة

**التقدير المجالي :**

إن تقدير المدى هو عبارة عن تقدير قيمة الثابت الإحصائي ضمن مجال محدد عند احتمال معين بدلالة التابع الإحصائي المقابل له آخذين بالحسبان الخطأ المعياري للتابع الإحصائي المراد التقدير بدلالته ،فنحصل على قيمة للثابت الإحصائي تقع عند ذلك الاحتمال بين حدين أعلى وأدنى ،وذلك بافتراض أن العينة العشوائية تتوزع توزيعا طبيعيا إذ يطلق على هذين الحدين اسم حدي الثقة أو مجال الثقة أو فترة الثقة،وعلى احتمال وقوع التقدير بين هذين الحدين باحتمال الدقة (درجة الثقة) والمتمم الحسابي لاحتمال الدقة بمستوى الدلالة ،فإذا رمزنا ب(α) لمستوى الدلالة فإن درجة الثقة يساوي (1-α) ،التي تعبر وتدل على المساحة المحصورة تحت المنحني الطبيعي

**القاعدة الأساسية للتقدير الإحصائي :**

**سوف نتطرق إلى التقدير المجالي :**

* **الأوساط الحسابية**
* **الانحرافات المعيارية**
* **النسب المئوية**
* **الفروق**

**تقدير فترة الثقة للأوساط الحسابية :**

لنفترض أن لدينا مجتمعا إحصائيا حجمه N وسحبنا منه جميع العينات العشوائية المتساوية الحجم فإنها تشكل فيما بينها توزيعا طبيعيا ,وكان الوسط الحسابي لهذا المجتمع غير معروف أما انحرافه المعياري σ فهو معروف ،ولنفترض أننا أخذنا عينة محددة من تلك العينات المسحوبة وحسبنا وسطها الحسابي

للمجتمع كالآتي **:**

**حدا الثقة = التابع الإحصائي الخطأ المعياري للتابع الإحصائي**

وعلية نكتب

وبشكل عام قد تكون قيمة الانحراف المعياري σ للمجتمع غير معروفة ، وفي هذه الحالة يمكننا الحصول على حدود الثقة السابقة باستخدام التقدير بدلالة الانحراف المعياري المحسوب من العينة شرط أن يكون حجمها أي

أي

**ملاحظة :** عندما يكون حجم العينة صغيرا والانحراف المعياري للمجتمع الإحصائي مجهولا فإنه بالإمكان اعتبار الانحراف المعياري المحسوب من العينة العشوائية تقديرا للانحراف المعياري للمجتمع الأصلي (بحسب قانون الأعداد الكبيرة ) إلا أن هذا التقدير يعد تقديرا متحيزا (ذا قيمة أقل من قيمة الانحراف المعياري للمجتمع الأصلي ) وبالتالي يجب تصحيح هذا التقدير المتحيز وذلك بضرب الانحراف المعياري للعينة بالعامل المصحح الذي يساوي وإن هذا التصحيح ضروري جدا عندما يكون حجم العينة صغيرا وبشكل عام :

S : الانحراف المعياري المتحيز للعينة صغيرة الحجم

**تقدير فترة الثقة للانحرافات المعيارية :**

اعتمادا على ما سبق :

**حدا الثقة = التابع الإحصائي \* الخطأ المعياري للتابع الإحصائي**

أي أن حدود الثقة للانحراف المعياريلمجتمع يتوزع توزيعا **،**كماهي مقدرة من عينة عشوائية سحبت منه وانحرافها المعياري ،تعطى كالآتي :

ولإيجاد حدود الثقة نستخدم كتقدير ل فيصبح :

**تقدير فترة الثقة للنسب المئوية :**

اعتمادا على توزيع معاينة النسب المئوية وعلى القاعدة الأساسية في إنشاء حدي الثقة حول قيمة التابع الإحصائي فإننا نضع حدود الثقة لنسب المجتمع كالآتي :

ونميز حالتين :

إذا كانت (P) معلومة :

إذا كانت (P) غير معلومة نستعيض ب

مثال : في استطلاع للرأي العام للعينة سحبت عينة عشوائية حجمها (100)من الناخبين في حي معين حيث دلت أن (55% ) منهم في صالح مرشح معين . فأوجد (99%) حدود الثقة للنسبة بين جميع الناخبين المؤيدين لهذا المرشح

الحل :

( 84.67 , 16.42)

أي باحتمال قدره (99%) فإننا واثقون بأن النسبة بين جميع الناخبين المؤيدين لهذا المرشح لن تقل عن 16%.42 ولن تزيد عن% 84.67

**تقدير فترة الثقة للفروق :**

إذا أخذنا عينتين عشوائيتين يقترب توزيع معاينتهما من التوزيع الطبيعي للمجتمع الإحصائي الذي أخذنا منه ،فإنه يمكننا تقدير فترة الثقة لثوابت المجتمع الإحصائي بدلالة تلك العينات ونميز الآتي :   
**تقدير فترة الثقة للفرق بين الأوساط الحسابية :**

حيث أن :

وإذا كانت تباينات المجتمعات الإحصائية غير معلومة يستعاض عنها بدلالة تباينات العينات (العينات مستقلة ) فتصبح :

**تقدير فترة الثقة بين الانحرافات المعيارية :**

وهنا لدينا حالتين :

* **معلومتين :**
* **غير معلومتين نحسبها بدلالة :**

**تقدير فترة الثقة لفروق النسب المئوية :**

1. ( معلومتين ):
2. ( غير معلومتين فتحسب بدلالة العينة ):

**اختبار الفرضيات :[[30]](#footnote-30)**

الفرضية الإحصائي هي تحديد قيمة ثابت إحصائي أي تعيين خاصة من خواص التوزيع الاحتمالي للمتغير العشوائي

كما نعلم أن توزيع المنحني الطبيعي وسطاؤه :

وحتى يكون التوزيع طبيعي معياري : يجب أن يكون قيمة الوسط الحسابي يساوي الصفر وقيمة الانحراف المعياري تساوي الواحد

أما في توزيع ثنائي الحدين : يتميز بمعرفة نسبة تحقق الصفة في المجتمع (P) حيث:

الوسط الحسابي :

*الانحراف المعياري :*

*لذلك لا بد لنا من معرفة طبيعة البيانات الإحصائية المتوفرة فهذا يساعدنا في اختيار أسلوب الاختبار الملائم لتلك البيانات*

* *إذا كانت البيانات كمية تخضع للتوزيع الطبيعي*
* *إذا كانت البيانات نوعية تخضع لتوزيع ثنائي الحدين*

*يعتبر توزيع ثنائي الحدين مقتربا من توزيع المنحني الطبيعي فيما*

*1- إذا كان حجم العينة أكبر أو يساوي 100 مفردة*

*2- إذا اقتربت النسبة المئوية لحدوث الصفة في العينة من النسبة المئوية لعدم حدوث الصفة في العينة ويصبح كل منها يساوي 50% نعتبر حينئذ أن توزيع المنحني الطبيعي مرضيا لتوزيع ثنائي الحدين*

عند القيام باختبار إحصائي لدينا فرضان

**الفرض الأول :** هو ما يسمى بفرض العدم وهي الفرضية التي تصاغ على أمل رفضها أي بمعنى لا يوجد اختلاف حقيقي أو جوهري بين التابع الإحصائي للعينة والثابت الإحصائي للمجتمع فيما إذا كانت المقارنة بين تابع وثابت (مجتمع موحد ) أما إذا كانت المقارنة بين تابع إحصائي آخر (اختبار يتعلق بمجتمعين ) وتنص الفرضية H0 عندئذ على أنه لا يوجد اختلاف جوهري بينهما

**الفرض الثاني :** يسمى بالفرضية البديلة H1 وهي على خلاف فرضية العدم H0 أي عند رفض الفرضية H0 فإننا نقبل الفرضية H1 أي يوجد اختلاف جوهري

**أنواع الأخطاء**

**خطأ من النوع الأول** : يحدث هذا النوع من الأخطاء عندما نقوم برفض فرضية العدم H0 ((رفض فرضية صحيحة )) واحتمال ارتكاب هذا الخطأ يدعى بمستوى الدلالة ويحدد مسبقا ويعبر عن احتمال الرفض الخاطئ أي رفض فرضية صحيحة

**خطأ من النوع الثاني** : يقع مثل هذا الخطأ عندما نقبل الفرض العدمي H0 في حين أنه خطأ وذلك باحتمال مقداره β ((قبول فرضية العدم ))

|  |  |  |
| --- | --- | --- |
| نتيجة الاختبار  حقيقة الفرضية | قبول الفرضية | رفض الفرضية |
| صحيحة | قرار سليم واحتماله(1-α) | رفض خاطئ احتماله (α)  خطأ من النوع الأول |
| خاطئة | قبول خاطئ احتماله (β)  خطأ من النوع الثاني | قرار سليم واحتماله (1-β)  ويسمى بقوة الاختبار |

**مستوى الدلالة** : وهو الحد الأقصى للخطأ (من النوع الأول ) المسموح بارتكابه وبالتالي طالما أن العينة عشوائية فهناك أخطاء احتمالية

مستوى الدلالة صغير احتمال الوقوع في الخطأ من النوع الثاني كبير

مستوى الدلالة كبير احتمال الوقوع في الخطأ من النوع الثاني صغير

لإجراء أي اختبار عن فرضية إحصائية ما ( عينات كبيرة الحجم أو صغيرة الحجم ) يراعى إتباع الخطوات التالية :

1. تحديد الفرضية : سواء كانت

فرضية عدم () :

**(فيما إذا كانت المقارنة بين تابع إحصائي وثابت إحصائي مقابل له )**

لا يوجد اختلاف حقيقي أو جوهري بين التابع الإحصائي والثابت الإحصائي المقابل له

**(أما إذا كانت المقارنة بين تابع إحصائي وتابع إحصائي مقابل له )**

نقول لا يوجد اختلاف حقيقي أو جوهري بين التابع الإحصائي والتابع الإحصائي المقابل له

أو فرضية بديلة : يوجد اختلاف حقيقي بين

التابع والثابت الإحصائي المقابل له (مقارنة تابع مع ثابت )

أو التابع الإحصائي والتابع المقابل له (مقارنة تابع مع تابع )

1. تحديد مستوى الدلالة (α) أو درجة الثقة أو احتمال الدقة (1-α) : ويتم تحديد مستوى الدلالة (α) اعتمادا على مدى درجة خطورة وفداحة الوقوع في الأخطاء من النوع الأول والأخطاء من النوع الثاني
2. تحديد قيمة النظرية المقابلة لمستوى الدلالة المعطى ، التي تؤخذ من جدول المساحات لتوزيع المنحني الطبيعي ، أي تحديد مناطق القبول أو الرفض في ضوء ذلك المستوى

الاختبار من اتجاهين :( ) ويكون مستوى الدلالة موزع على طرفي التوزيع بالتساوي

الاختبار من اتجاه واحد : ( ) وفيها يقع مستوى الدلالة كاملا على طرف واحد من التوزيع

* قيمة التابع الإحصائي أكبر أو أقل من قيمة الثابت الإحصائي ،كان الاختبار من اتجاهين
* قيمة التابع الإحصائي أكبر أو أفضل أو أحسن من قيمة الثابت الإحصائي كان الاختبار من اتجاه واحد ويقع على الطرف الأيمن من التوزيع
* قيمة التابع الإحصائي أقل أو أسوأ من قيمة الثابت الإحصائي كان الاختبار من اتجاه واحد ويقع على الطرف الأيسر من التوزيع

1. تحديد قيمة الاختبار والمحسوبة في ضوء البيانات المعلومة (تختلف باختلاف المؤشر المراد اختباره )
2. مجتمع إحصائي واحد :
3. مجتمعين إحصائيين :
4. المقارنة اتخاذ القرار الإحصائي من خلال مقارنة قيمة أو مع قيمة الاختبار المحسوبة ،فالقرار الإحصائي يقسم مجال قيمة الاختبار المحسوبة إلى منطقتين ،تسمى إحداهما بمنطقة القبول والأخرى بمنطقة الرفض . ومنطقة القبول هي المنطقة التي تتكون من قيمة الاختبار المحسوبة وكثيرة الحدوث إذا كانت صحيحة ، أما منطقة الرفض فهي المنطقة التي تتكون من قيمة الاختبار قليلة الحدوث إذا كانت صحيحة وبناء على ذلك يكون القرار برفض إذا وقعت قيمة الاختبار في منطقة الرفض

* :نقبل الفرضية البديلة ونرفض فرضية العدم
* نقبل فرضية العدم ونرفض الفرض البديل

يقسم القرار الإحصائي قيمة الاختبار (Z) إلى منطقتين :

1. منطقة القبول : وهي المنطقة التي تتكون من قيمة الاختبار (Z) وكثيرة الحدوث إذا كانت صحيحة .
2. منطقة الرفض : وهي المنطقة التي تتكون من قيمة الاختبار (Z) وقليلة الحدوث إذا كانت صحيحة

بناء على ما سبق نميز ثلاث حالات :

* الفرضية ، والفرضية البديلة

في هذه الحالة نميز حالتين إما :

:

أي أننا أمام اختبار من اتجاهين

* الفرضية ، والفرضية البديلة

في هذه الحالة نضع (α) كاملة في الطرف الأيمن من توزيع قيمة الاختبار (Z)

* الفرضية ، والفرضية البديلة

أي أننا أمام اختبار من اتجاه واحد فقط ، في هذه الحالة نضع (α) كاملة على طرف واحد من قيمة الاختبار (Z) من الجهة اليسرى

**الاستدلال الإحصائي للعينات صغيرة الحجم**

**توزيع T ستودنت**

من المفترض استخدام توزيع ستودنت T كلما كان الانحراف المعياري للمجتمع الإحصائي غير معروف حيث يتم تقديره من الانحراف المعياري للعينة أي كلما كان الخطأ المعياري ليس حقيقيا بل مقدرا .

إلا أن منحنى Z يستخدم كتقريب جيد وممتاز لمنحنيات T

وبشكل عام

1. إذا كان لدينا مجتمع ما له متوسط وانحراف معياري σ سحبت منه عينة كبيرة الحجم n ،لها وسط حسابيX وانحراف معياري S فإن X سوف يتبع التوزيع الطبيعي بمتوسط قدره ự وبانحراف معياري وبالتالي فالمتغير Z يساوي

(وذلك عند و معلوم )

1. أما إذا كان الانحراف المعياري للمجتمع σ غير معلوم فيمكن استبداله بالانحراف للعينة s عندها سيظل المتغير Z يأخذ الشكل
2. أما عندما تكون 30> n وكان المجتمع الإحصائي الذي سحبت منه العينة المدروسة يتوزع توزيعا طبيعيا له وسط حسابي وانحراف معياري σ معلوم فإن متوسط العينة سوف يتبع التوزيع الطبيعي بوسط حسابي ،وانحراف معياري سيظل المتغير يأخذ الشكل :
3. أما عندما30 >n وكان المجتمع الإحصائي الذي سحبت منه العينة المدروسة يتوزع توزيعا طبيعيا بوسط حسابي انحراف معياري σ مجهول فإننا سوف نستبدل σ المجهول بالانحراف المعياري للعينة S ففي هذه الحالة الأخيرة سوف يأخذ المتغير توزيعا جديدا هو توزيع T ستودنت ،وذلك بدرجات حرية قدرها طالما كان الانحراف المعياري للمجتمع σ مجهولا ،وحجم العينة صغيرا حيث

وبصورة أخرى : فإن الأساليب الإحصائية المعتمدة في التحليل الإحصائي للعينات صغيرة الحجم ،هي نفسها تلك الأساليب التي تم اعتمادها في التحليل الإحصائي للعينات كبيرة الحجم باستثناء استبدال المتغير Z بالمتغير t عند درجات حرية ومستوى دلالة و  
يرمز له .

**أهمية توزيع ستودنت T**

إن أهمية توزيع ستودنت تنبع من إمكانية استعماله في حالات العينات صغيرة الحجم وهو وضع واقعي كثير الحدوث حيث أن المحددات العملية من ضيق للإمكانيات وعدم توفر المعدات والآلات المطلوبة لأخذ العينات تجعل تصغير حجم العينة أمرا ضروريا

**تعريف (T) النظرية :**

وتسمى أيضا الجدولية أو الحرجة وهي عبارة عن الحد الأقصى المسموح بارتكابه من الأخطاء علما أن هذه الأخطاء من النوع الأول ولإيجاد هذه القيمة نحتاج إلى :

* مستوى الدلالة
* حجم العينة : لأن كل حجم له توزيع خاص ومنحني خاص ، أما في العينات كبيرة الحجم (التوزيع الطبيعي ) فإن شكل المنحني واحد .

**جدول T واستخدامه**

يتواجد توزيع T الآن في معظم البرامج الإحصائية الجاهزة شأنه في ذلك شأن التوزيع الطبيعي المعياري ولحساب الاحتمالات تحت منحني دالة الكثافة الاحتمالية للمتغير T والمحصورة بين أي قيمتين من قيم المتغير يلزمنا حساب التكامل لهذه الدالة التكاملية وإن لكل قيمة من قيم منحني خاصا بها . ويتم حساب المساحات في هذا الجدول استنادا إلى درجات الحرية ومستوى الدلالة المعتمد النظرية (الجدولية ).فجدول توزيع ستودنت مؤلف من عناوين للأعمدة وهي تعبر عن مستوى الدلالة ( ) وهي المساحات للطرف الأيمن الموجب لهذا المنحني (جدول من طرف واحد ) . أما العمود الأول فهو يحتوي على درجات الحرية

مثال : احسب قيمة

أي احسب قيمة t النظرية عند و

**مفهوم درجات الحرية :**

تعرف درجات الحرية بأنها عدد الدرجات أو عدد التكرارات التي يمكن أن تتغير حول قيمة ثابتة أو مقياس معين للمجتمع الإحصائي أو هي حجم العينة مطروحا منه عدد الثوابت الإحصائية ،والتي يجب تقديرها من مفردات العينة

**خصائص توزيع T ستودنت:**

من أهم صفات وخصائص ستودنت t ما يلي :

1. يمكن التعرف على توزيع t وشكله (هناك عدد لا نهائي من التوزيعات t ) بمعرفة درجات الحرية فقط .
2. المنحني الاحتمالي للتوزيع له قمة واحدة يشبه الجرس وهو متماثل حول الصفر وبالتالي فمتوسط ذلك التوزيع يساوي الصفر و .
3. تباين توزيع t يساوي والتباين دائما أكبر من الواحد الصحيح . وبالتالي عند زيادة درجات الحرية أو زيادة حجم العينة يقل التباين ،ويقترب من الواحد الصحيح ،وبالتالي يصبح أكثر تجانسا وأكثر اقترابا مع التوزيع الطبيعي
4. إن دالة الكثافة الاحتمالية لتوزيع t لا يمكن الاستفادة منها عمليا فالمساحة تحت المنحني تم حسابها في جداول خاصة وذلك للدراسات التطبيقية والعملية لمعظم المشاكل الإحصائية المتطلبة لهذا التوزيع .
5. بالمقارنة مع المنحني الطبيعي فهو أكثر تفرطحا ،بمعنى أن المساحة في طرفيه أكبر في توزيع t من التوزيع الطبيعي
6. يتطابق توزيع ستودنت t مع التوزيع الطبيعي عند درجات حرية كبيرة لأن توزيع ستودنت t مع ازدياد درجات الحرية يزيد ارتفاع منحناه ويصبح أقل تشتتا وأكثر تدببا وبالتالي ينطبق وبشكل كبير مع المنحني الطبيعي

**تطبيقات واستخدامات توزيع ستودنت**

عندما تكون العينة صغيرة الحجم وتباين المجتمع الذي سحبت منه هذه العينة مجهولا يستخدم توزيع t في :

1. تقدير فترة الثقة للمتوسطات .
2. الاختبارات الإحصائية للمتوسطات .

**تقدير فترة الثقة للمتوسطات**

**تقدير فترة الثقة لمتوسط المجتمع ( )** : إن فترة الثقة تمثل منطقة قبول الفرضيات الإحصائية وخارج هذه الفترة تمثل منطقة الرفض للفرضيات الإحصائية وبصورة أخرى نجد أنه توجد علاقة وثيقة بين اختبار الفرضيات وفترة الثقة فهما أسلوبان لموضوع واحد هو الاستدلال الإحصائي

بالإمكان استنتاج فترة الثقة لمتوسط المجتمع من الاختبار الإحصائي وذلك باستبدال قيمة t النظرية مكان t المحسوبة ، حيث t النظرية هي النقطة التي تفصل بين قبول الفرضيات ورفض الفرضيات الإحصائية والاختبار الإحصائي لتوزيع t :

وبعد استبدال t النظرية مكان t المحسوبة نجد :

*حدا الثقة لمتوسط المجتمع (فترة الثقة )*

حيث أن و

مثال : إذا كانت الإنتاجية اليومية لعينة عشوائية من سبعة عمال من عمال أحد المصانع كالتالي :

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| 7 | 6 | 5 | 4 | 3 | 2 | 1 | العامل |
| 14 | 14 | 10 | 11 | 16 | 9 | 10 | الإنتاجية (بالقطعة ) |

احسب فترة الثقة باحتمال 95% للإنتاجية اليومية لجميع عمال هذا المصنع بفرض أن الإنتاجية اليومية تتوزع وفق التوزيع الطبيعي أو قريب من الطبيعي .

الحل :

درجات الحرية

النظرية أو الجدولية

وبالتالي فترة الثقة لمتوسط إنتاجية جميع عمال هذا المصنع هي (95 %)

) إنتاجية جميع عمال هذا المصنع الذي سحبت منه العينة المدروسة لن يزيد عن 95%باحتمال قدره (

14,447 قطعة ولن يقل عن9,553

**تقدير فترة الثقة لقيمة الفرق الحقيقي بين الوسطين الحسابيين للمجتمعين لعينتين مستقلتين بافتراض تساوي التباين بينهما**

**ملاحظات :**

حجوم العينات المستقلة إذا كانت متساوية فليس هناك داعي لدراسة المجتمع ونعتبرها متجانسة

عندما تكون العينات مستقلة فإن

إذا وجد من العينتين المستقلتين أنهما متساويتا الحجم فهذا يعني أن التباين في المجتمعين الأصليين متساويين وليس من حاجة لدراسة التباينات عن طريق توزيع فيشر والعكس صحيح .

**تقدير فترة الثقة لقيمة الفارق الحقيقي بين الوسطين الحسابين للمجتمعين لعينتين مستقلتين تباينهما غير متساو**

ملاحظات هامة :

عندما نقول عينات مستقلة فيكون : إما حجم العينات متساوي

أو حجم العينات غير متساوي

عندما نقول عينات غير مستقلة : حتما حجم العينات يكون متساوي

إذا كانت التباينات متساوية نلجأ إلى التباين المشترك

إذا كانت التباينات غير متساوية نلجأ إلى درجة الحرية التقريبية

لا نستطيع الاختبار عن طريق حدا الثقة إلا إذا كان الاختبار من اتجاهين

**تقدير فترة الثقة لقيمة الفرق الحقيقي بين الوسطين الحسابيين لمجتمعين من عينتين غير مستقلتين**

توزيع فيشر :

يعد توزيع فيشر من أهم التوزيعات المستخدمة في دراسة الإحصاء التطبيقي ومن أهم استخدامات هذا التوزيع :

اختبارات الفرضيات الإحصائية لتساوي التباينين لمجتمعين مستقلين

اختبارات الفرضيات الإحصائية لتساوي أكثر من متوسطين وهو ما يسمى بتحليل التباين (ANOVA)

وبالتالي F دائما موجبة

**قائمة المراجع:**

1. **عيد,يحيى.بحوث التسويق والتصدير.القاهرة:مطابع سجل العرب.ط1. 1996.**
2. **سكران,أوما.ت:بسيوني,اسماعيل علي.طرق البحث في الإدارة.الرياض:دار المريخ,2006.**
3. **عبيدات,محمد .وآخرون.منهجية البحث العلمي.عمان:دار وائل للنشر.ط2. 1999.**
4. **بري,عدنان.وآخرون.مبادئ الاحصاء والاحتمالات.الرياض:مطابع جامعة الملك سعود.ط1 .1991.**
5. **محمد,أمانى موسى.التحليل الاحصائي للبيانات.القاهرة:مركز تطوير الدراسات العليا.2007.**
6. **تشاو,لنكولن.ت:عزام,عبد المرضي.الاحصاء في الادارة.الرياض:دار المريخ.2004.**
7. **شامل,محمد بهاء الدين.الإحصاء بلا معاناة.الرياض:مركز البحوث.2006.**
8. **القاضي,دلال.وآخرون.الإحصاء للإداريين والاقتصاديين.عمان:دار الحامد.2005.  
   حميدان ,عدنان عباس . مخول ,مطانيوس . جاعوني ,فريد . ناصر آغا , عمار .الإحصاء التطبيقي : دمشق :منشورات جامعة دمشق .2006**
9. **حميدان ,عدنان عباس . ناصر آغا , عمار .الجاعوني ,فريد .العواد ,منذر .مبادئ الإحصاء : دمشق :منشورات جامعة دمشق . 2004**

1. يحيى عيد.بحوث التسويق والتصدير.القاهرة:مطابع سجل العرب.ط1. 1996.ص306-307. [↑](#footnote-ref-1)
2. أوما سكران.ت:اسماعيل علي بسيوني.طرق البحث في الإدارة.الرياض:دار المريخ,2006.ص438. [↑](#footnote-ref-2)
3. محمد عبيدات,وآخرون.منهجية البحث العلمي.عمان:دار وائل للنشر.ط2. 1999.ص116. [↑](#footnote-ref-3)
4. محمد عبيدات,وآخرون.المرجع السابق مباشرة.ص117-118. [↑](#footnote-ref-4)
5. محمد بهاء الدين شامل.الإحصاء بلا معاناة.الرياض:مركز البحوث.2006.ص176. [↑](#footnote-ref-5)
6. دلال القاضي, وآخرون.الإحصاء للإداريين والاقتصاديين.عمان:دار الحامد.2005.ص53. [↑](#footnote-ref-6)
7. لنكولن تشاو,ت:عبد المرضي عزام.الاحصاء في الادارة.الرياض:دار المريخ.2004. [↑](#footnote-ref-7)
8. أمانى موسى محمد.التحليل الاحصائي للبيانات.القاهرة:مركز تطوير الدراسات العليا.2007.ص33. [↑](#footnote-ref-8)
9. دلال القاضي, وآخرون.مرجع سبق ذكره.ص63. [↑](#footnote-ref-9)
10. لنكولن تشاو, ت:عبد المرضي عزام.مرجع سبق ذكره. [↑](#footnote-ref-10)
11. عدنان بري.وآخرون.مبادئ الاحصاء والاحتمالات.الرياض:مطابع جامعة الملك سعود.ط1 .1991.ص78. [↑](#footnote-ref-11)
12. محمد عبيدات, وآخرون.مرجع سبق ذكره.ص125. [↑](#footnote-ref-12)
13. دلال القاضي.وآخرون.مرجع سبق ذكره .ص70. [↑](#footnote-ref-13)
14. محمد بهاء الدين شامل.مرجع سبق ذكره.ص181. [↑](#footnote-ref-14)
15. عدنان بري, وآخرون.مرجع سبق ذكره.ص82 [↑](#footnote-ref-15)
16. عدنان بري.وآخرون.المرجع السابق مباشرة.ص83-84 [↑](#footnote-ref-16)
17. دلال القاضي,وآخرون.مرجع سبق ذكره.ص74. [↑](#footnote-ref-17)
18. محمد بهاء الدين شامل.مرجع سبق ذكره.ص185. [↑](#footnote-ref-18)
19. عدنان بري,وآخرون.مرجع سبق ذكره.ص93 [↑](#footnote-ref-19)
20. عدنان بري, وآخرون.مرجع سبق ذكره.ص94 [↑](#footnote-ref-20)
21. محمد بهاء الدين شامل,.مرجع سبق ذكره.ص187. [↑](#footnote-ref-21)
22. محمد بهاء الدين شامل.مرجع السابق مباشرة.ص187-188. [↑](#footnote-ref-22)
23. امانى موسى محمد.مرجع سبق ذكره [↑](#footnote-ref-23)
24. لنكولن تشاو.ت: عبد المرضي عزام.مرجع سبق ذكره. [↑](#footnote-ref-24)
25. دلال القاضي,وآخرون.مرجع سبق ذكره.ص92 [↑](#footnote-ref-25)
26. دلال القاضي,وآخرون.المرجع السابق مباشرة.ص93-94 [↑](#footnote-ref-26)
27. أمانى موسى محمد.مرجع سبق ذكره.ص59 [↑](#footnote-ref-27)
28. . ,عدنان عباس حميدان . عمار ناصر آغا , . فريد الجاعوني , . منذر العواد , .مبادئ الإحصاء : دمشق :منشورات جامعة دمشق . 2004 [↑](#footnote-ref-28)
29. . عدنان عباس حميدان , مطانيوس مخول , فريد جاعوني , عمار ناصر آغا .الإحصاء التطبيقي : دمشق :منشورات جامعة دمشق .2006 [↑](#footnote-ref-29)
30. عدنان عباس حميدان وآخرون .الإحصاء التطبيقي : دمشق :منشورات جامعة دمشق .2006 [↑](#footnote-ref-30)