مبادئ نظرية العينات

١-١ مقدمة

تم نظرية العينات بدراسة العلاقة بـ ا تمع والعينات ا سحوبة منـه فيمـا يـسمى بالاسـتدلال الإحصائي ، statistical inference. هناك عدة طرق لأخذ العينات من ا تمع لاستخدامها الاستدلال الإحصائي ومن أشهر هذه الطرق هي العينة العشوائية وهي العينة الـ تكون لكل مفرده من مفردات ا تمـع نفس فرصة الاختيار العيينة. فمثلاً نستع بعينه مسحوبة من ا تمع لتقدير معا هذا ا تمع مثل متوسـطه أو تباينة أو غ ذلك. أو أعطاء عينه من ا رضى بارتفاع الضغط، مثلاً دواء مع قياس ضغطهم قبل وبعـد تناو م ذا الدواء عرفة ما إذا كان هذا الدواء مفيد خفض الضغط أم لا.

أي موعات من ا فردات تشترك صفه أصفات وتكون موضوع دراسة أو ث فإن هذه ا موعة يطلق عليها إحصائيا تمع الدراسة (أو اختصاراً ا تمع Population). وا تمع قد يكون موعة مـا مـن البشر أو أشجار أنواع معينه من الفاكهة أو ا يوانات الزراعية أو إنتاج دولة ما لسلع معينه خلال فترة زمنية دده...ا .

وا تمع قد يكون دوداً إذا كان كن حصر عدد أفراده مثل سكان مدينة ما أو طـلاب مرحلـة تعليمية معينة ، وقد يكون ا تمع غ دود (لا ائي) إذا كان لا كن حصر عـدد أفـراده مثـل النجـوم والكواكب أو الكائنات ا ية ياه ا يطات والأ ار وعند دراسة صفة ما أو صفات معينه تمع مـا ، فـإن البيانات الإحصائية عن تلك الصفة أو الصفات مع بأحد أسلوب

أولاً: أسلوب ا صر الشامل (census) وفيه مع البيانات عن كل مفرده من مفـردات ا تمـع، وهـذا الأسلوب يتطلب وفرة الوقت وا ال وا هود الف وتزداد هذه ا تطلبات وتتضاعف كلمـا ازداد حجـم ا تمع (عدد أفراد ا تمع). وهذا الأسلوب لا يتبع عادة إلا حالة التعدادات الـ ريها الـدول وتـدعمها بإمكانيات ضخمه مثل تعدادات السكان والتعدادات الصناعية والتعدادات الزراعية.

الثا : أسلوب ا عاينة (Sampling method) وفيه يتم ع البيانات عن جزء من مفردات ا تمـع تـار بطريقة أو بأخرى ويطلق عليه عينه (Sample) بعد ذلك يتم تعميم نتائج الدراسة على ا تمع بأكمله. أي أن إسلوب ا عينة يقصد به دراسة خصائص ا تمع من خلال دراسة عينه مسحوبة منه ، و اح هذا الإسـلوب يعتمد على أن مل العينة أقصى درجة من دقة التمثيل للمجتمع ا سحوبة منه.

من الأفضل بعض ا الات ا صول على معلومات دقيقة عن طريق التعداد التام أو ا صر الشامل ميـع عناصر ا تمع، لكن لاستخدام أسلوب ا عاينة فوائد ة مقارنة بالتعداد الشامل يرد بيا ا الفقرة التالية.

١-٢ بعض مزايا أسلوب ا عاينة

يتميز أسلوب ا عاينة عن أسلوب ا صر الشامل زايا عدديه منها:

١- يؤدي استخدام العينات العشوائية إ خفض تكاليف الدراسات ا يدانية بسبب صغر حجم العينة بالنسبة إ حجم ا تمع وهو ما يؤدي إ فيض الأعباء الإدارية والفنية ال تتطلبها أي دراسة ميدانية.

٢- يتحقق وفر واضح الوقت الذي ينفق دراسة ميدانية على أساس عينة بدلاً من ا صر الشامل وتتضح أ ية الوقت عندما نقوم بدراسة ظاهرة تتغ رور الوقت ، وعينة قد يترتب على دراسة تلك الظاهرة ا تمع كله مع البيانات من يع مفردات ا تمع أن ر وقت بديل فتكون البيانات والنتـائج وقـت ظهورها غ مطابقة لواقع ا تمع وتصبح النتائج ذات قيمه دودة بعد أن فقدت عنصر ا طابقة مع واقع الظاهرة وتوزيعها ا ا تمع ، والتعدادات الدورية للسكان وبسبب ضخامة حجم العمل ا تـستغرق وقتاً طويلاً ح تصبح نتائجها جاهزة ومنشورة وقد يطول هذا الوقت إ أكثر من ثلاث أو أربع سنوات ح مع استخدام أحدث أجهزة ا اسبات الأليه الضخمة ، ويكون على الباحث مستخدمي هذه النتائج مراعاة الوقت الذي ينقض ب تاريخ إجراء التعداد وتاريخ نشر نتائجه وتعديل هذه النتائج حدود ذلك .. وهذا دفع الكث من الدول إ تعزيز نتائج التعدادات الدورية للسكان بنتائج تعدادات ري ب كل تعدادين متتالي على أساس العينة.

٣- ا تمعات غ ا دودة (اللا ائية) مثل تمع الكائنات ا ية البحار وا يطـات لا كـن أن تـتم الدراسة على أساس ا صر الشامل ولكن لابد وأن تتم الدراسة بأسلوب ا عاينة.

٤- أيضاً هناك بعض الاختبارات لابد وأن تتم بأسلوب ا عاينة لأن إجراء مثل هذه الاختبارات على أسـاس ا صر الشامل يؤدي إ تلف ا ادة ا ة أو هلاكها.. فاختبار صلاحية شحنه من ا فرقعات مثلاً لابـد وأن يتم على أساس العينة وبا ثل ليل دم ا رضى يتم على أساس عينه.

١-٣ أقسام العينات

تنقسم العينات عادة إ قسم رئيس و ا عينات عشوائية وعينات غ عـشوائية، وفيمـا يلـي تفصيل لكل قسم منها.

١-٣-١ العينات العشوائية

هي تلك العينات ال يتم اختيار مفردا ا حسب خطه إحصائية لا يكون فيها للباحث أو فـردات العينة دخل اختيار أي مفرده فيها ، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلا ا الدور الأول اختيار ا فردة ولكن بشرط أن يتحقق يع ا فردات احتمال ثابت و ـدد للاختيـار. والعينـات العشوائية إذا ما اختيارها بالطريقة العلمية السليمة وا ناسبة كن أن تكفل درجه عالية من دقـة التمثيـل

للمجتمعات ا سحوبة منها لذلك فهي الوسيلة الأساسية حالة البحوث العلمية الدقيقة .. من أهـم أنـواع العينات العشوائية مايلي.

(أ) العينة العشوائية البسيطة: Simple random sample

ويلجأ إليها الباحث حالة ما إذا كان تمع الدراسة ليس كب اً و مل قدراً من التجـانس بـ ا فردات للصفة أو الصفات موضع الدراسة. والعينة العشوائية البسيطة تستغل فرص متكافئة فردات ا تمع للدخول العينة ولكن ا فردات ال تدخل العينة تكون عن طريق الصدفة البحتة. والاختيار العشوائي يتم يدوياً عن طريق بطاقات متماثلة ا جم واللون أو عن طريق جداول الأعداد العـشوائية أو عـن طريـق ا اسب الآ . ولكي يتحقق ذلك فإن الأمر يتطلب ديد مفردات ا تمع ديداً كاملاً ويكون هذا التحديـد على شكل قائمة (أو خريطة) تضم كل مفردات ا تمع وهذه القائمة تسمى الإطـار (Frame) ولا ـوز الاختيار العشوائي إلا من ا فردات ال يضمها الإطار.

(ب) العينة ا نتظمة: Systematic sample

اختيار هذه العينة يتطلب وجود إطار للمجتمع كما حالة العينة العشوائية البسيطة يـث يعطـى لكل مفرده من مفردات ا تمع رقماً متسلسلاً داخل الإطار، تار مفردات العينة من الإطار يث يكـون الرقم ا تسلسل لكل مفردة يبعد بعداً ثابتاً منتظماً عن رقم ا فردة السابقة ا وكذلك رقم ا فردة اللاحقة ا. فمثلاً إذا كان لدينا تمعاً حجمه 2000 مفرده ونريد اختيار عينه منتظمة حجمها 100 مفرده فإننـا نقـسم الإطار إ فترات منتظمة طول كل فترة $= \frac{2000}{100} = 20$ مفردة ومن داخل مفردات الفترة الأو (1 – 20) تار مفرده واحدة عشوائياً ولتكن رقم 14 مثلاً وبناء على رقم تلك ا فردة يتحدد بـاقي مفـردات العينـة ا نتظمة فتكون هي ا فردات ذات الأرقام 34 ، 54 ، .... ، 1974، 1994.

والعينة ا نتظمة كث ة الاستعمال التطبيقات العملية لقلة تكاليفها وقلة الأخطاء الـ ترتكـب اختيـار مفردات العينة فضلاً عن سهولة إجرائها. ولكن أهم عيوب ا عاينة ا نتظمة هو عدم صلاحيتها إذا ما وجـدت علاقة دورية مع ترتيب العناصر القائمة وكان طول الفترة بـ عناصر العينة مساوياً لطول الدورة أو إحدى مضاعفا ا.

(جـ) العينة العشوائية الطبقية: Stratified random sample

ويلجأ إليها الباحث حالة ما إذا كان تمع الدراسة واضحاً به فئات (طبقات) يث أن التجانس أو التقارب داخل كل طبقة من طبقات تمع الدراسة أكـ من التجانس داخل ا تمع ككل (أي أن التـشتت داخل ا تمع ككل أكـ من التشتت داخل كل فئة من فئاته على حده). هذه ا الة ـب علـى الباحـث مراعاة أن الطبقة داخل العينة بنفس نسبة وجودها داخل ا تمع (وأحياناً يوضع الاعتبار عناصر أخرى مثل التشتت داخل الطبقة أو عنصر التكلفة مع البيانات عن الطبقة). بعد أن يتم ديد عدد ا فردات الـ ـب

سحبها من كل طبقة للدخول العينة فإن هذه ا فردات يتم سحبها عشوائياً من داخل الطبقة و موع هـذه ا فردات تكون العينة الطبقية العشوائية.

(د) العينة متعددة ا راحل أو العنقودية: clustered sample

يلجأ إليها الباحث عندما يكون تمع الدراسة كب جداً ومتناثراً على مساحات شاسعة تكلف الكث من الوقت وا هد التنقل بينها عند ع البيانات، أيضاً حالة عدم وجود إطار يضم يع مفردات ا تمع فيستحيل الاختيار العشوائي مباشر من ا تمع. ذا يلجأ الباحث إ أخذ العينة على مراحل متعددة متتالية. ا رحلة الأو يتم تقسيم ا تمع إ عدد دد من وحدات ا عاينة الكب ة ا جم ومنها تار بعضها عشوائياً يتلو ذلك كمرحلة ثانية تقسيم الوحدات ا ختارة عشوائياً من ا رحلة الأو إ وحدات أقل منها ا جم تار بعضها عشوائياً.. وهكذا تتابع مراحل التقسيم والاختيار العشوائي ، وعدد هذه ا راحل ليس ثابت بـل يتوقف على طبيعة تمع الدراسة وإمكانيات الباحث .. ا رحلة الأخ ة يصل الباحث إ وحدات ا عاينـة ال سيجمع عنها بيانات البحث ويطلق عليها وحدات ا عاينة الأولية.

١-٣-٢ العينات غ العشوائية:

هي تلك العينات ال لا تكفل ميع مفردات ا تمع احتمال ثابت و دد للاختيار، وغالباً يتـدخل الباحث عملية الاختيار بصورة أو بأخرى ... ومن أهم أنواع العينات غ العشوائية:

(أ) العينة العمدية أو ا قصودة: Purposive sample

يلجأ الباحث إ هذه الطريقة فيما إذا كان تمع الدراسة كب جداً وكانت إمكانياته لا تسمح لـه إلا بدراسة عينة حجمها صغ جداً بالنسبة تمع الدراسة، هذه ا الة يتعمد الباحث اختيار مفردات معينـة كعينة تمع الدراسة يرى ته السابقة أن هذه العينة كن أن تعطي ثيلاً مقبولاً تمع الدراسة.

مثلاً إذا أراد باحث دراسة خصائص اقتصادية أو اجتماعية معينه عن ريف دولـة مـا ، وكانـت إمكانياته ا الية والإدارية لا تسمح له بعينة سوى سكان قرية واحدة ، فإنه هذه ا الة إذا ما اختيار القرية عشوائياً من ب آلاف القرى بتلك الدولة فإن الصدفة قد تأ بقرية بعيدة خصائصها (من حيـث الظـاهرة موضوع الدراسة) عن خصائص معظم قرى تلك الدولة ... كأن تأ بالصدفة قرية ساحلية معظم سكا ا مـن الصيادين أو قرية قريبة من مشروع صناعي ضخم يستوعب فواه العاملة معظم سكا ا.. هذه القرية أو تلك قد يأخذ النمط ا عيشي لسكناها طابعاً خاصاً – نابعاً عن ظروفها ا اصة – بعيداً عن النمط ا عيـشي ا عتـاد لبقية القرى، لذلك فأي منها لا كن أن يعطي ثيلاً مقبولا لريف تلك الدولة. ذا فإن الباحث وعلى ضـوء خ اته السابقة يتعمد اختيار قرية معينة يرى أ ا – من وجهة نظره الشخصية- كن أن ثل الريـف. وهـذه الطريقة غ علميه وغالباً يتم اللجوء إليها حالة البحوث التمهيدية.

(ب) العينة الحصصية: Quota sample

وهي نوع خاص من العينات غير العشوائية وتستخدم كثيراً في معاينة الرأي العام (على سبيل المثال عمليات استطلاعية الرأي العام التي يقوم بها معهد جالوب قبل إجراء انتخابات الرئاسة في الولايات المتحدة الأمريكية)... في هذه الطريقة يقسم المجتمع موضوع الدراسة إلى طبقات بالنسبة إلى صفات أو خصائص معينة ويتم العمل على تمثيل كل طبقة منها في العينة بنسبة وجودها في المجتمع الأصلي (وعلى سبيل المثال في حالة دراسة الدخل لمنطقة ما ورؤى أن يكون حجم العينة المطلوبة ١٠٠ فرد مثلاً عندما يريد الباحث أن يقوم جامعوا البيانات بالحصول على البيانات من ٢٠ موظفاً، ٤٥ من العمال في الريف، ٣٥ من ذوي الأعمال الحرة.. وتترك الحرية لجامعي البيانات في اختيار الأفراد المطلوبة فيها حدود المواصفات الموضوعة لكل طبقة من الطبقات المذكورة.

واضح أنه رغماً من أن هذه الطريقة في ظاهرها مماثلة للعينة الطبقية العشوائية.. إلا أنه في الحالة الأخيرة (العينة الطبقية العشوائية) يكون اختيار المفردات عشوائياً من داخل كل طبقة ولا يترك لجامع البيانات حرية اختيار المفردات من كل طبقة والذي قد يترتب عليه تحيزاً كبيراً.

عموماً.. يلجأ الباحث إلى العينة الحصصية إذا كان من المرغوب فيها اظهار النتائج في وقت قصير مع التغاضي عن توافر درجة دقة عالية بتلك النتائج.

١-٤ أخطاء البيانات الإحصائية

تتعرض البيانات الإحصائية التي يتم جمعها إلى نوعين من الأخطاء:

١-٤-١ خطأ التميز

وهو ينتج عن مصادر متعددة، منها أخطاء في تصميم البحث أو التجربة أو أخطاء فنيه أثناء جمع البيانات أو خلال العمليات الحسابية التي تتم على البيانات المجمعة.. أخطاء التميز تزداد بازدياد الفروق بين الإمكانيات (المادية والفنية) اللازم توافرها لضمان أقصى درجة دقة ممكنه وبين الإمكانيات الفعلية المتاحة للباحث.

أخطاء التميز قد توجد في البيانات التي يتم جمعها بأسلوب الحصر الشامل وقد توجد أيضاً في البيانات التي يتم جمعها بأسلوب المعاينة، ولكنها إن وجدت فهي غالباً أكبر في الحالة الأولى (الحصر الشامل) مما هي عليه في الحالة الثانية (المعاينة) باعتبار أن حجم العمل في تلك الحالة يكون أقل وبالتالي قد يسهل توفير الإمكانيات اللازمة وتجنب الأخطاء الفنية.

١-٤-٢ خطأ ا عاينة العشوائية أو خطأ الصدفة

وهو ا طأ الناتج عن فروق الصدفة ب مفردات ا تمع ا دخلت العينة و تلك ا فردات ا تـ شأ الصدفة أن تدخل العينة.. وخطأ الصدفة كن تقليل قيمته إذا ما اختيار العينة بالطريقة ا ناسبة وإذا ما كان حجم العينة مناسباً جم ا تمع وخصائصه.

١-٥ ا عا والإحصاءات

ا قاييس الإحصائية ا سب من بيانات تمع الدراسة بأكملـه يطلـق عليهـا معـا ا تمـع (Parameters of population)، أما ا قاييس الإحصائية ا سب من بيانات عينة مسحوبة من تمـع الدراسة فيطلق عليها إحصاءات (Statistics) ويعت كل إحصاء منها ثابة تقدير أو قيمه تقديريـة علمـة ا تمع ا ناظر، فيكون ا توسط ا سا ا سوب من بيانات العينة تقدير علمة ا تمع ا ناظرة وهي ا توسـط ا سا سحوب منه هذه العينة وهكذا.. و ب ألا يغيب عن الأذهان بأن حساب قيمة ا توسط ا سا من بيانات العينة ليس هدفاً حد ذاته ولكن وسيلة للتعرف على ا توسط ا سا للمجتمع موضوع الدراسـة.. وهكذا با ال بالنسبة لباقي ا قاييس الإحصائية ا سب من العينة.

للتفرقة ب ا عا والإحصاءات ب أن نرمز لكل منها برموز تلف عن رموز الأخرى، على سـبيل ا ثال يرمز للمتوسط ا سا للمجتمع بالرمز μ بينما يرمز للمتوسط ا سا للعينـة بـالرمز $\bar{x}$ ، أيـضاً للا راف ا عياري للمجتمع بالرمز σ بينما يرمز للا راف ا عياري للعينة بالرمز S وهكذا.

١-٥-١ توزيعات ا عاينة: Sampling Distributions

نفرض أننا أخذنا عينه حجمها n من تمع ما ، سحبنا منها بعض ا قاييس الإحصائية مثل ا توسط ا سا ، التباين ، ... فإن كل مقياس من هذه ا قاييس يعت متغ عشوائي ذاته تلف من عينة إ أخرى – هذا ا تغ العشوائي ضع لتوزيع مع – هذا التوزيع يسمى بتوزيع العينة. فمثلاً نقول أن توزيع ا عايـنة للمتوسط ا سا وهو عبارة عن توزيع يع ا توسطات ا سابية للعينات ا أخوذة من نفس هذا ا تمع ذات ا جم n ، وكذلك فإن توزيع ا عاينة للتباين هو توزيع يع التباينات ا سوبة من عينات ا نفس ا جـم n ومأخوذة من نفس ا تمع ، وهكذا ... .

١-٥-٢ توزيعات ا عاينة للأوساط: Sampling Distributions of Means

نفرض أننا سحبنا عينه حجمها n من تمع لا نهائي ، القيمة ا توقعة له تساوي μ والا راف ا عياري هو σ فإن ا توسط ا سا $\bar{X}$ ضع لتوزيع ما ، متوسط هذا التوزيع وا راف ا عياري ا

$$\sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}} \quad ، \quad \mu_{\bar{X}} = \mu \qquad (4\text{-}1)$$

و ا الة ال يكون فيها ا تمع الأصلي ا سحوبة منه العينة تمع طبيعي (ويرمز له بالرمز $N(\mu, \sigma^2)$ فإن توزيع ا عاينة للمتوسط ا سا $\bar{X}$ يكون هذه ا الة توزيع طبيعي أيضاً له نفس ا توسط الأصلي μ ولكن ا راف ا عياري يساوي $\sigma/\sqrt{n}$ ، أي ع أن

$$X \sim N(\mu, \sigma^2) \quad \Rightarrow \quad \bar{X} \sim N(\mu, \frac{\sigma^2}{n}) \qquad (4\text{-}2)$$

ومن يكون

$$z = \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \sim N(0,1) \qquad (4\text{-}3)$$

أما إذا كان ا تمع غ طبيعي فإن $\bar{X}$ لا ضع للتوزيع الطبيعي ولكنها تتوزع توزيع يكون قريباً من التوزيع الطبيعي لقيم n الكب ة $(n \geq 30)$ حيث أن

$$z = \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \xrightarrow[n \to \infty]{as} N(0,1) \qquad (4\text{-}4)$$

وتعت النتيجة السابقة ا امة جداً الإحصاء وخاصة التطبيقات العلمية وتسمى نظرية النهاية ا ركزية Central Limit Theorem وال تنص على أنه حالة العينات الكب ة ا جم فإن ا توسط ا سا $\bar{X}$ ضع للتوزيع الطبيعي با عاملات μ و $\frac{\sigma^2}{n}$ ، حيث أن σ^2, μ ا متوسط وتباين ا تمع الأصلي بغض النظر عن شكل توزيع ا تمع الأصلي. ومن فإنه لقيم n الكب ة تتحقق العلاقة (4-3) بصرف النظر عن توزيع ا تمع الأصلي.

كذلك فإنه إذا كان $\bar{X}_1$ هو ا توسط ا سا لعينه عشوائية مسحوبة من تمع لا ائي متوسطه هو μ_1 وا راف ا عياري هو σ_1 ، وكان $\bar{X}_2$ هو ا توسط ا سا لعينة عشوائية مسحوبة من تمع لا ائي آخر متوسط μ_2 وا راف ا عياري σ_2 وكانت العينت مستقلت فإن ا موع ا ي توسط العينت ضع لتوزيع ا عاينة با عاملات

$$\mu_{(\bar{x} \pm \bar{x}_2)} = \mu_1 \pm \mu_2 \quad \text{and} \quad \sigma^2_{(\bar{x}_1 \pm \bar{x}_2)} = \frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2} \qquad (4\text{-}5)$$

حيث n_1, n_2 ا حجم العينة الأو والثانية.

وإذا كان ا تمع الأصلي طبيعي فإن $(\bar{x}_1 \pm \bar{x}_2)$ ضع لتوزيع طبيعي أيضاً بالبارامترات ا عطاة (4-5) وعليه فإنه هذه ا الة

$$z = \frac{(\bar{x}_1 \pm \bar{x}_2) - (\mu_1 \pm \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1) \qquad (4\text{-}6)$$

أما إذا كان أحد ا تمع أو كليهما لا يتوزع توزيعاً طبيعياً فإن $(\bar{x}_1 \pm \bar{x}_2)$ لا يتوزع توزيعاً طبيعياً كذلك ، ولكن لقيم n_1, n_2 الكب ة فإنه طبقاً لنظرية النهاية ا ركزية السابقة فإن $(\bar{x}_1 \pm \bar{x}_2)$ يتوزع توزيعاً قريباً من التوزيع الطبيعي وبذلك كننا استخدام نفس العلاقة (4-6) حالة العينات الكب ة.

١-٥-٣ توزيع ا عاينة للتباين: Sampling Distribution of The Variance

إذا كان $s^2 = \frac{\sum (x_i - \bar{x})^2}{n-1}$ هو تباين عينه عشوائية حجمها n مأخوذة من تمع متوسطه μ وتباينه σ^2 وعزمه الرابع حول ا توسط هو μ_4 فإن

$$\mu_{s^2} = \sigma^2 \qquad \text{and} \qquad \sigma^2{}_{s^2} = \frac{\mu_4 - \sigma^2}{n-1} \tag{4-7}$$

وإذا كان ا تمع طبيعي فإن $\mu_4 = 3\sigma^2$ وبالتالي فإن

$$\sigma^2_{s^2} = \left(\frac{2}{n-1}\right)\sigma^2 \tag{4-8}$$

نلاحظ هنا أن S^2 لا تتوزع طبيعي حـ ولو كان ا تمع طبيعي ، ولكنه يتوزع توزيع قريب من التوزيع الطبيعي وذلك لقيم n الكبـ ة $(n \geq 100)$. أما إن كان ا تمع الأصلي ضع للتوزيع الطبيعي فإن ا تغ ة $(n-1)s^2/\sigma^2$ ضع لتوزيع يسمى توزيع مربع كاي χ^2 بعدد درجات حرية يساوي $n-1$. أي أن

$$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2(n-1) \tag{4-9}$$

ويعتبر توزيع مربع كاي من التوزيعات ا امة الإحصاء التطبيقي ودالة كثافته هي

$$f(y) = y^{\frac{v-1}{2}} e^{-y/2} \quad , \quad y > 0 \tag{4-10}$$

حيث v هي عدد درجات ا رية للتوزيع وتعت هي ا عامل الوحيد له ويتضح من شكل الدالة أ ا داله متصلة وتقع بأكملها فوق النصف ا وجب ور السينات ، منح هذه الداله غ متماثل وبعت من ا نحنيات موجبة الالتواء ويقل التواؤه (وبالتالي يقترب من التماثل) كما زادت درجات ا ريه v . وتكون القيمة ا توقعة لهذا التوزيع هي v و تباينه هو $2v$ أي أن

$$E(y) = \mu_y = v$$

$$V(y) = \sigma^2 = 2v \tag{4-11}$$

فإذا كان s_1^2 هو تباين عينه عشوائية حجمها n_1 مسحوبة من تمع طبيعي $N(\mu_1, \sigma_1^2)$ ، وكان s_2^2 هو تباين عينه عشوائية أخرى حجمها n_2 ومسحوبة من تمع طبيعي آخر $N(\mu_2, \sigma_2^2)$ وكانت العينتان مستقلتان فإن ا تغ ة :

$$\frac{s_1^2/\sigma_1^2}{s_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1) \tag{4-12}$$

حيث أن $F(n_1-1, n_2-1)$ تسمى بتوزيع F بدرج ا ريه $n_1 - 1$ و $n_2 - 1$ و دالة الكثافة الإحتماليه للمتغ y الذي ضع لتوزيع F بدرج ا ريه v_1, v_2 تعطى بالصورة:

$$f(y) = \frac{y^{\frac{v_1}{2}-1}}{(v_1 y + v_2)^{\frac{v_1+v_2}{2}}} \quad , \quad y > 0 \qquad (4\text{-}13)$$

وكما يتضح من الداله (4-13) أن ا نحـ يقع بالكامل النصف ا وجب ور السينات كمـا حالـة توزيع χ^2 ، وهو أيضاً غ متماثل وموجب الالتواء ولكن يقترب من التماثل كلما زادت درجـات ا ريـه v_1, v_2 .

ذكرنا سابقاً أنه إذا كان $\bar{X}$ هو ا توسط ا سا لعينه حجمها n مأخوذة من تمع طبيعي با عاملات μ, σ^2 فإن

$$z = \frac{\sqrt{n}(\bar{x} - \mu)}{\sigma} \quad \sim \quad N(0,1)$$

هذا إذا كانت σ معلومة ، ولكن حالة ما إذا كانت قيمة σ غ معلومة فإننا نستخدم بدلا منها الا راف ا عياري للعينة S ولكن هذه ا الة يصبح ا تغ $\frac{\sqrt{n}(\bar{x}-\mu)}{S}$ ضع لتوزيع يعرف بتوزيع t ستيودنت $t-$ student بدرجات حريه $n-1$ ، أي أن

$$t = \frac{\sqrt{n}(\bar{x} - \mu)}{s} \quad \sim \quad t(n-1) \qquad (4\text{-}14)$$

دالة الكثافة لتوزيع t بدرجات حريه v تعطي بالصورة:

$$f(t) = \left(1 + \frac{t^2}{v}\right)^{-\frac{v+1}{2}} \qquad (4\text{-}15)$$

وهو توزيع متماثل حول ور y وهو يشبه ذلك ا نحـ الطبيعي القياسي $N(0,1)$ ولكنه أقل دباً مـن التوزيع الطبيعي القياسي ولكنه يقترب من التوزيع الطبيعي كلما زادت درجات ا ريه.

وإذا كان $\bar{X}_2$ و S_1^2 ا ا توسط ا سا والتباين لعينه حجمها n_1 مأخوذة من تمع طبيعي متوسط هو μ_1 وكان $\bar{X}_2$ و S_2^2 ا ا توسط ا سا والتباين لعينه أخرى حجمها n_2 ومأخوذة من تمع طبيعي آخر لـه ا توسط μ_2 وكانت العينتان مستقلتان فإن ا تغ

$$t = \frac{(\bar{X}_2 \pm \bar{X}_2) - (\mu_1 \pm \mu_2)}{S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad \sim \quad t(n_1 + n_2 - 2) \qquad (4\text{-}16)$$

حيث أن $S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$ **يسمى بالتباين ا شترك للعينة** The Pooled Variance.