الفصل الحادي عشر

اختبارات العينه الواحدة المتضمنة للمتوسطات والتناسب

Single - Sample Tests Involving Means & Proportions

سوف نركز في هذا الفصل على اختبارات فروض عن المتوسطات والتناسب الخاصة بمجتمع البحث · وسوف نقارن بين متوسط أو نسبة نحصل عليها من عينة واحدة وبين مَعلمٍ مفترض ، (Hypothesized Parameter) ، وقرار اتخذ برفض أو عدم رفض الفرض . وسوف نكتشف بسرعة أن نوع الاختبارات التي تعرض في هذا الفصل لها فائدة عملية أقل من الفائدة المرجوة من الاختبارات الخاصة بالعينات المتعددة . وعلى كل حال فإن الأمر الأكثر أهمية هنا هو الحصول علي فهم جيد للأفكار الأساسية وليس على التطبيقات العلمية . ومن سوء الحظ فإن الاختبارات الأبسط ليست دائمًا الأكثر فائدة .

عند إعداد الاختبارات الإحصائية المتضمنة نظرية ذات الحدين تكون لدينا المقدرة لنرى بالضبط كيف تستخدم نظرية الاحتمالات للحصول على توزيع المعاينة Sampling Distribution ومن هنا وفي الفصول التالية سوف تصبح الإعتبارات الرياضية أكثر تعقيداً لدرجة أننا يجب أن نأخذ كثيرًا منها كقضايا مسلمة مهما كانت أهمية فهم الخلفية الرياضية لأي مسألة · بالطبع توجد البراهين الرياضية لكل هذه القضايا ، ولكن معظمها يتضمن إجراءات ونظريات معقدة ومستويات متقدمة في نظرية الاحتمالات وخلفية متقدمة في الرياضيات .

١١- ١ توزيع المعاينة للمتوسطات :-

Sampling Distribution of Means.

هناك نظرية جديرة بالملاحظة مبنية على نفس مبادئ نظرية الاحتمالات وقواعدها كما هو الحال بالنسبة لنظرية ذات الحدين . ولكن لايمكن التطرق إلي إثباتها في مثل هذا الكتاب ، ويمكن كتابة هذه النظرية علي النحو التالي: إذا اخترنا عينات عشوائية متكررة ذات حجم (ن) من مجتمع بحث ذي توزيع طبيعي ، وبمتوسط حسابي س (μ) وتباين ع^{٢} (σ^2) فإن توزيع المعاينة سيكون توزيعاً طبيعياً (Normal Distribution) بمتوسط حسابي س وتباين $\frac{ع^{٢}}{ن}$ ($\frac{\sigma^2}{N}$) . دعنا نلقي الان نظرة فاحصة على ماتقول هذة النظرية .

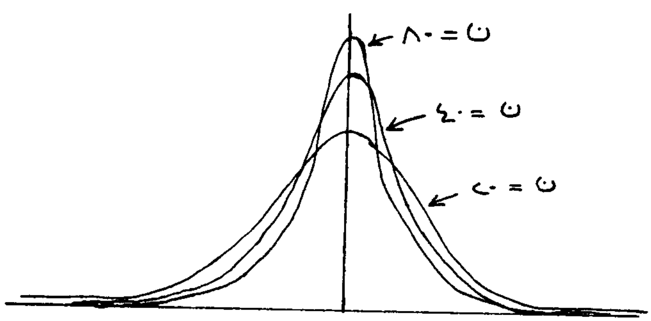
نبدأ أولاً بمجتمع ذي توزيع طبيعي (مع إدراكنا أنه لايوجد بالطبع في الواقع مايسمي بالمجتمع الطبيعي الكامل) ثم نتخيل بعد ذلك أننا نقوم بسحب عدد كبير من عينات عشوائية ذات حجم (ن) من هذا المجتمع [١] . ثم نحصل علي المتوسط ($\bar{س}$) ($\bar{x}$) لكل واحدة من هذه العينات . ومن الطبيعي أن نجد أن هذه المتوسطات سوف تختلف إلي حد ما من عينة لأخري . ولكننا نتوقع أن تختلف علي شكل عنقودي حول المتوسط الحقيقي لمجتمع البحث (س) . إن النظرية تقول أننا لو أعددنا رسماً بيانياً لتوزيع متوسطات العينات فإن الناتج سيكون منحني طبيعيـاً (Normal Curve) . علاوة علي ذلك ، فإن الانحراف المعياري للمنحني الطبيعي لمتوسطات

---

(١) لابد من الحذر من الخلط بين عدد العينات (غير المحدود) والحجم (ن) لكل عينة .

العينات سيكون $\frac{\text{ع}}{\sqrt{\text{ن}}}$ [$\frac{\sigma}{\sqrt{N}}$] ٠ وعليه فكلما كان حجم العينة المختارة كبيراً ، كان الإنحراف المعياري لتوزيع المعاينة صغيراً ٠ وهذا يعني زيادة التحلق علي الشكل العنقودي (أنظر الشكل ١١-١)٠ اذا اعتبرنا متوسطات العينات كتقديرات لمتوسط مجتمع البحث فإننا نستطيع القول أن هناك مقداراً مامن الخطأ في عملية التقدير راجع إلي تذبذب

مقارنة التوزيعات الطبيعية لأحجام مختلفة من العينات



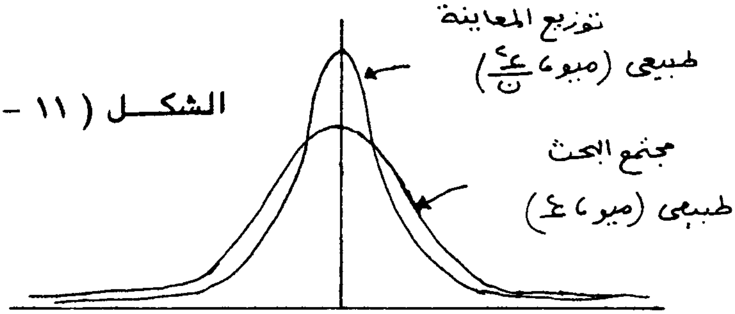
الشكل (١١-١)

المعاينات ٠ لهذا فإننا نسمي .الانحراف المعياري لتوزيع المعاينة باسم الخطأ المعياري (Standard Error) ٠ وفي هذه الحالة يصبح الخطأ المعياري للمتوسط الذي أشرنا إليه من قبل بالرمز $\text{ع}_{\bar{س}}$ ($\sigma_{\bar{x}}$) كالأتي:

$\frac{\text{ع}}{\sqrt{\text{ن}}}$ ($\frac{\sigma}{\sqrt{N}}$) ٠ ويجب أن نتذكر دائماً أن هناك ثلاثة توزيعات مميزة متضمنة، منها اثنان طبيعيان. أولا : مجتمع البحث

المفترض أن يكون موزعًا توزيعاً طبيعيًا ، بوسط حسابي س وتباين عـ ٢ (σ^2)] بعدئذ ستكتب خصائص هذا المجتمع بالصيغة المختصرة التالية : طبيعي (سَ ، عـ٢)] [NOR (μ , σ^2] . ثانيًا : هناك توزيع للدرجات داخل كل عينة . فإذا كان حجم العينة (ن) كبيرًا فإنّ هذا التوزيع سيكون ممثلاً لمجتمع البحث بشكل معقول ، وبالتالي يكون توزيعًا طبيعيًا تقريبيًا .لاحظ أن هذا هو التوزيع الوحيد الذي يمكن الحصول عليه تجريبيًا (٢) .ثالثًا هناك توزيع المعاينة للمؤشر الإحصائي (المتوسط الحسابي) الذي سيكون توزيعه طبيعيًا . لقد رأينا الأن أن توزيع المعاينة للمتوسط سوف يكون أيضًا له توزيع طبيعي . ولكن سيكون إنحرافه المعياري أصغر من الإنحراف المعياري لمجتمع البحث (مالم يكن حجم العينة (ن) يساوي واحدًا) .

إن العلاقة بين مجتمع البحث وتوزيع المعاينة تتضح من الشكل رقم (١١-٢) . وكلما كان حجم العينة (ن) كبيرًا ، فإن قمة توزيع المعاينة تكون أكثر حدة كما هو واضح في الشكل رقم (١١- ٢) أدناه . فيجب أن يكون



الشكل (١١ - ٢)

مقارنة بين توزيع مجتمع البحث وتوزيع المعاينة.

(٢) بما أن هذا هو التوزيع الذي يراه الباحث فعلاً فإن من الممكن أن يكون هنالك اتجاه لخلط هذا النوع الثاني من التوزيع بتوزيع المعاينة

واضحًا في أذهاننا ، ونتذكر دائمًا أنه بالرغم من أن الإنحرافات المعيارية للتوزيعات مرتبطة ببعضها بعضًا مباشرة ، إلا أن هذه التوزيعات مختلفة تمامًا. إن كل الحالات في توزيع المعاينة هي متوسطات لعينات منفصلة . كما كان الحال مع نظرية ذات الحدين ، وكما هو الحال في كل الاختبارات الاحصائية الأخري ، فإن توزيع المعاينة وليس مجتمع البحث هو الذي يستخدم مباشرة في اختبارات الدلالة . إن الفروض عن مجتمع البحث ربما تظهر في النموذج . إن الحكم عن مجتمع البحث وعن أساليب الحصول علي العينات تتحول بواسطة نظرية الاحتمال إلى أحكام توزيع المعاينة . وباختصار فإن المتوسطات والانحرافات المعيارية لأنواع التوزيعات الثلاثة تكون على النحو التالي :-

نوع التوزيعات	المتوسط الحسابي	الانحراف المعياري
مجتمع البحث	سَ (μ)	عـ (σ)
العينة	سَـَ ($\bar{X}$)	ع (S)
توزيع المعاينة	سَ (μ)	$\frac{\text{عـ}}{\sqrt{\text{ن}}}$ $\frac{\sigma}{\sqrt{N}}$

إن هذه النظرية تتفق مع الفهم العام الذي مفاده أن الشخص يثق بتقدير المتوسط من عينة كبيرة أفضل من ثقته في تقديره من عينه صغيرة،مع إفتراض أنه قد تم تجنب التحيًز [٣] . وفي الواقع فإن النظرية تقول إن الاختلاف بين متوسط عينة وأخري سوف يكون أقل كلما كان حجم العينة (ن) كبيرًا . ولكن النظرية تمثل تنقيحًا كبيرًا للفهم العام حيث تقدم

لنامؤشراً عن درجة الثقة التي لدينا إذا زاد حجم العينة بمقدار معين . فعلي سبيل المثال لكي نخفض حجم الخطأ المعياري (Standard Error) إلي النصف علينا أن نزيد حجم العينة إلي أربعة أضعاف . وكذلك توضح النظرية أنه كلما كان مجتمع البحث متجانساً منذ البداية ، قلت قيمة الانحراف المعياري وقل حجم الخطأ المعياري $\left(\frac{ع}{\sqrt{ن}}\right)$ ، وزاد تمركز متوسطات العينات حول متوسط مجتمع البحث . إن التبرير النظري لهذه النظرية الهامة يكون بإدخال فكرة التوافيق الخطية (Linear Combinations) ، وهي الفكرة التي سوف نستخدمها لاحقًا في عدة نقاط . إن المتوسط هو بالطبع دالة خطية بسيطة عن الدرجات $\bar{س}$ حيث إن : -

$$\bar{س} = \frac{1}{ن} \left[س_1 + س_2 + \cdots + س_ن \right]$$

وعموماً يمكن إثبات أنه لو كان لدينا االمتغير (ص) (y) الذي هو دالة خطية لقيم $س_i$ (x_i) ؛ وإذا تم اختيار قيم $س_و$ بطريقة مستقلة ، كما هو الحال عندما نسحب عينة عشوائية فيمكننا الحصول علي تعبيرات بسيطة عن المتوسط [القيمة المتوقعة] لقيم المتغير (ص) ، وكذلك عن التباين لقيم المتغير (ص) . وبالتحديد إذا كان : -

---

(٣) لاحظ أن لدينا ثقة أكبر في التقديرات التي تنبني على عينات كبيرة ولكن عند رفض الفرض الصفري عند مستوي دلالة ٠٥.و ، فإننا نخاطر مخاطرة الخطأ نفسها من النمط الأول بصرف النظر عن حجم العينة (ن) . وكما سنري لاحقاً فإن حجم المنطقة الحرجةالمستخدمة في الاختبار يأخذ حجم العينة في الاعتبار . وهكذا نتحسب للتضارب الظاهر.

$$ص = جـ_١ سـ_١ + جـ_٢ سـ_٢ + جـ_٣ سـ_٣ + \cdots\cdots + جـ_ن سـ_ن$$

$$Y = C_1 X_1 + C_2 X_2 + C_3 X_3 + \ldots\ldots C_N X_N$$

وإذا تم اختيار قيم المتغير سـ [سـ و] اختياراً مستقلاً يكـــون لدينــا الأتي :-

$$توقع (ص) = جـ_١ توقع(سـ_١) + جـ_٢ توقع (سـ_٢) + \cdots + جـ_ن توقع سـ_ن$$

$$E(y) = C_1 E(X_1) + C_2 E(X_2) + \ldots\ldots C_N E(X_N)$$

$$وتباين ص = عـ ص^٢ = جـ_١^٢ عـ س_١^٢ + جـ_٢^٢ عـ س_٢^٢ + \cdots + جـ_ن^٢ عـ س_ن^٢$$

$$Var(Y) = \sigma_y^2 = C_1^2 \sigma_{X_1}^2 + C_2^2 \sigma_{X_2}^2 + \ldots\ldots C_N^2 \sigma_{X_N}^2$$

(١) في حالة العينات العشوائية فإن القيمة المتوقعة لكلٍّ من [سـ و] هي $\bar{س}$ (μ) . وإذا اعتبرنا أن جـ و = $\frac{١}{ن}$ ، [$C_i = \frac{1}{N}$] ، فإن (ص) تصبح متوسط العينة ويكون لدينا : -

$$توقع (\bar{سـ}) = توقع (ص) = \frac{١}{ن} [\bar{س} + \bar{س} + \cdots + \bar{س}] =$$

$$\frac{١}{ن} (ن \bar{س})] = \bar{س}$$

$$E(\bar{X}) = E(Y) = \frac{1}{N} [\mu + \mu + \ldots\ldots \mu] = \frac{1}{N} [N\mu] = \mu$$

$$وحيث إنّ جـ_١^٢ = جـ_٢^٢ = \cdots\cdots جـ_ن^٢ = \frac{١}{ن^٢}$$

$$\text{فإنّ } \text{عـ}_{\text{ص}}^{2} = \text{عـ}_{\bar{\text{س}}}^{2} = \frac{1}{\text{ن}^{2}}\,\text{عـ}_{\text{س}_1}^{2} + \frac{1}{\text{ن}^{2}}\,\text{عـ}_{\text{س}_2}^{2} + \cdots + \frac{1}{\text{ن}^{2}}\,\text{عـ}_{\text{س}_{\text{ن}}}^{2}$$

$$= \frac{1}{\text{ن}^{2}} \left[\text{عـ}^{2} + \text{عـ}^{2} + \cdots + \text{عـ}^{2}\right] = \frac{1}{\text{ن}^{2}} \left(\text{ن}\ \text{عـ}^{2}\right) = \frac{\text{عـ}^{2}}{\text{ن}}$$

$$\sigma Y^2 = \sigma X^2 = \frac{1}{N^2}\,\sigma X_1^2 + \frac{1}{N^2}\,\sigma X_2^2 + \ldots \frac{1}{N^2}\,\sigma X_N^2$$

$$= \frac{1}{N^2}\left[\sigma^2 + \sigma^2 + \ldots \sigma^2\right] = \frac{1}{N^2}\left(N\sigma^2\right) = \frac{\sigma^2}{N}$$

وعليه نكون قد أوضحنا أن توقع $(\bar{\text{س}})$ = سَ ، وأن $\text{عـ}_{\bar{\text{س}}}^{2} = \frac{\text{عـ}^{2}}{\text{ن}}$

وعليه يكون $\text{عـ}_{\bar{\text{س}}} = \frac{\text{عـ}}{\sqrt{\text{ن}}}$

إنّ نتيجة الخطأ المعياري هذه تنبع من حقيقة أن تباين كلٌ من $\text{س}_{\text{و}}$ هو مجرد عـ² (σ^2) ، حيث أننا نتعامل مع حالات فردية تم إختيارها باحتمالات متساوية من مجتمع بحث له تباين عـ² (σ^2) . والفكرة بداهة هي أننا لو كررنا تجربة المفردة الأولي عدداً كبيراً جداً من المرات فإن توزيع هذه المرات سيكون توزيعاً طبيعياً تقريباً بمتوسط حسابي سَ وتبايـــن عـ² $Nor\ (u, \sigma^2)$ ونفس الشيء سيكون صحيحاً بالنسبة لحساب المفردة الثانية وهكذا .

نظرية النهاية المركزية : The Central - Limit Theory

نحن الآن في وضع يسمح لنا بالتطرق إلي نظرية أكثر عمومية تعرف باسم نظرية النهاية المركزية • نصها كالأتي :

إذا تكرر سحب عينات عشوائية من مجتمع بحث ، حجم كلٍّ منها (ن) [مهما كان شكل مجتمع البحث] ، وإذا كان لهذا المجتمع متوسط حسابي ($\bar{س}$)وتباين ع^{٢} ، فإنه كلما أصبح حجم ا لعينة كبيراً ، إتَّجه بالتالي توزيع المعاينة لمتوسطات العينات نحو التوزيع الطبيعي مع متوسط حسابي $\bar{س}$ ،

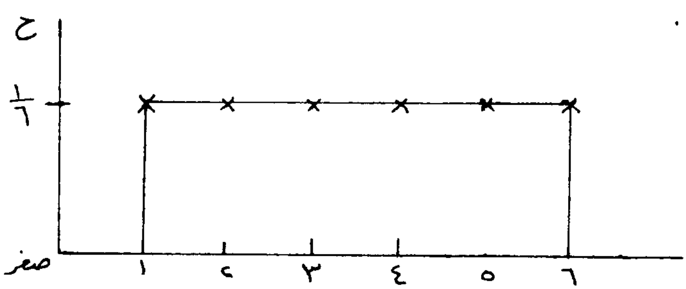
وتباين $\frac{ع^{٢}}{ن}$.

إن هذه النظرية جديرة بالملاحظة أكثر من النظرية السابقة .هذه النظرية تقول إنه مهما كان التوزيع الذي بدأنا به غير عادي (على شرط أن يكون الحجم (ن) كبيراً بما فيه الكفاية) يمكن أن نعتمد في النهاية علي توزيع معاينة في شكل طبيعي تقريباً • وحيث أن توزيع المعاينة (وليس توزيع مجتمع البحث) هو الذي سيستخدم في إختبارات الدلالة ؛ فإنَّ هذا يعني أنه كلما كانت(ن) كبيرة استطعنا أن نطمئن كلية على إفتراض التوزيع الطبيعي لمجتمع البحث ، ونستفيد من المنحني الطبيعي في إختباراتنا .

إننا نحتاج فعلاً إلي أدلة مقنعة تؤكد أن نظرية النهاية المركزية معقولة تجريبياً • وإنَّ أفضل وسيلة للحصول علي فكرة جيدة عن معني نظرية النهاية المركزية ، وفي نفس الوقت لاقناع الباحث نفسه بأن الخطأ المعياري هو حقيقة $\frac{ع}{\sqrt{ن}}$ [$\frac{\sigma}{\sqrt{N}}$] ، هي أن تأخذ عدداً من العينات من مجتمع بحث له متوسط حسابي معلوم وانحراف معياري معلوم ثم نحسب متوسط العينات ثم نوجد الانحراف المعياري لهذه المتوسطات ونقارن النتيجة مع

المعادلة $\frac{ع}{\sqrt{ن}}$ [$\frac{\sigma}{\sqrt{N}}$].[٤] ولكن لماذا يصبح توزيع المعاينة طبيعياً إذا كان توزيع مجتمع البحث ليس طبيعياً ؟ دعنا نلقي نظرة علي مجتمع بحث يبعد عن التوزيع الطبيعي ثم نري ماذا يحدث عندما نأخذ عينات تتفاوت في الكبر. تخيل أننا نقذف بزهرة نرد (غير متحيزة من الناحية الرياضية) بحيث يكون إحتمال الحصول علي كل وجه من الوجوه الستة يساوي $\frac{١}{٦}$ بالضبط . إن توزيع الإحتمالات لقذف زهرة نرد واحدة يكون مستطيلا.

أي أن كلّ الأرقام من واحد إلي ستة لها فرص حدوث متساوية . إن هذا النوع من التوزيعات في تناقض واضح مع التوزيع الطبيعي الذي تكون فيه القيم المتطرفة أقل إحتمالاً من القيم القريبة من المتوسط (إن هذا التوزيع المستطيل يمكن تمثيله بالشكل رقم (١١ - ٣) . وإذا أردنا الدقة فإن التوزيع سيكون منقطعاً ولن يكون مستمراً كما يوحي بذلك الشكل رقم (١١-٣) .



الشكل (١١- ٣)

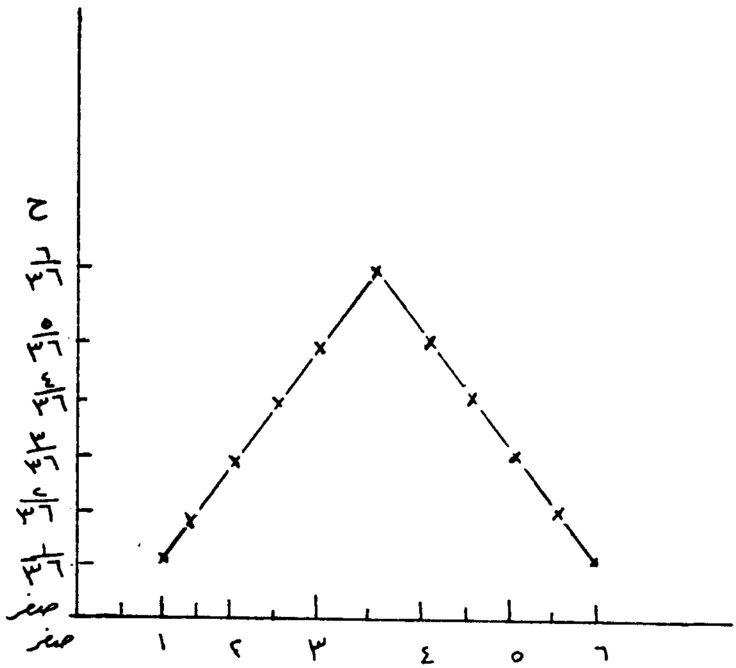
باعتبار مثل هذا التوزيع لمجتمع البحث مكونًا من كلّ المرات الممكنة لقذف زهرة النرد ، دعنا نحسب توزيع المعاينة لمتوسطات العينات التي حجم كل منها يساوي ٢. وهذا يعني أننا سوف نقذف زهرتي نرد ثم نجمع قيم

(٤)انظر المسألة رقم (١) في نهاية الفصل .

الزهرتين وبعد ذلك نقسم المجموع على إثنين · وكما يعي جيداً لاعبو الطاولة المتمرسون فإن مدي هذا المجموع يتراوح بين ٢ و ١٢ مع اعتبار أن المجموع الذي يساوي ٧ هو الأكثر إحتمالاً · وللحصول علي احتمالات حدوث كل من حالات الجمع هذه نلاحظ أن هناك ٦×٦ أي ٣٦ ناتجاً محتملاً إذا ميزنا بين زهرتي النرد . فالأولى يمكن أن تبرز أي وجه من وجوهها الستة وكذلك زهرة النرد الثانية · وللحصول علي مجموع «٧» من الزهرتين وبالتالي متوسط ٣ر٥ ، فإننا نحتاج لحساب عدد الطرق التي يمكن أن نحصل بواسطتها على هذه النتيجة . ومن الواضح أن هناك ستة أزواج يمكن أن تعطينا مجموع «٧» وهي (١ ، ٦) ، (٢ ، ٥) ، (٣ ، ٤) ،(٤ ، ٣) ، (٥ ، ٢) و (٦ ، ١)·أما مجموع «٦» فيمكن الحصول عليه بخمس طرق فقط وهى (١ ، ٥) ، (٢ ، ٤) ، (٣ ، ٣) ، (٤ ، ٢) ، (٥ ، ١) . كما أن هناك حالة واحدة فقط نحصل فيها على مجموع «١٢» وهى (٦ ، ٦) ، وحالة واحدة أيضًا نحصل فيها على مجموع «٢» وهى (١ ، ١). يمكننا إذن أن نمثل التوزيع الاحتمالي للمتوسطات علي النحو التالي : –

المتوسط	الاحتمال	المتوسط	الاحتمال
١,٠	$\frac{١}{٣٦}$	٤,٠	$\frac{٥}{٣٦}$
١,٥	$\frac{٢}{٣٦}$	٤,٥	$\frac{٤}{٣٦}$
٢,٠	$\frac{٣}{٣٦}$	٥,٠	$\frac{٣}{٣٦}$
٢,٥	$\frac{٤}{٣٦}$	٥,٥	$\frac{٢}{٣٦}$
٣,٠	$\frac{٥}{٣٦}$	٦,٠	$\frac{١}{٣٦}$
٣,٥	$\frac{٦}{٣٦}$		
		المجموع	$\frac{٣٦}{٣٦}$

اذا مثلنا قيم هذه الاحتمالات برسم بياني ، فإن توزيع المعاينة سيأخذ شكل مثلث كما هو موضح في الشكل رقم (١١ - ٤) .



الشكل (١١-٤)

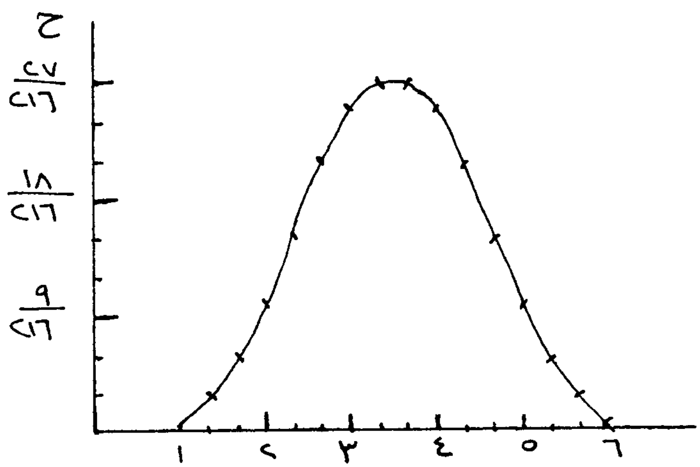
توزيع المعاينة لمتوسطات الحصول على أرقام زهرتي نرد غير متحيزتين، عند رميهما

إذا قذفنا بثلاث قطع من زهرة النرد وجمعنا الأوجه الثلاثة وحصلنا علي المتوسطات فإن توزيع المعاينة للمتوسطات سيكون علي النحو التالي :-

المتوسط	الاحتمال	المتوسط	الاحتمال
١,٠٠	١/٢١٦	٣,٦٧	٢٧/٢١٦
١,٣٣	٣/٢١٦	٤,٠٠	٢٥/٢١٦
١,٦٧	٦/٢١٦	٤,٣٣	٢١/٢١٦
٢,٠٠	١٠/٢١٦	٤,٦٧	١٥/٢١٦
٢,٣٣	١٥/٢١٦	٥,٠٠	١٠/٢١٦
٢,٦٧	٢١/٢١٦	٥,٣٣	٦/٢١٦
٣,٠٠	٢٥/٢١٦	٥,٦٧	٣/٢١٦
٣,٣٣	٢٧/٢١٦	٦,٠٠	١/٢١٦
		المجموع	٢١٦/٢١٦

إن هذا التوزيع كما يظهر ببساطة في الشكل رقم (١١- ٥) بدأ يأخذ شكل المنحني الطبيعي ، بالرغم من أن العينة حجمها ثلاث مفردات فقط . بعد إلقاء نظرة فاحصة على الأرقام الواردة في الجدول أعلاه يجب أن يكون الشخص قادرًا علي أن يدرك بالحد س ماذا يحدث ، ولماذا يأخذ المنحني شكل الجرس تدريجيًا كلما كبر حجم العينة ن . وبالرغم من أن احتمال الحصول علي الرقم ٦ في حالة قذف زهرة نرد هو نفس احتمال الحصول علي الرقم ٣ أو الرقم ٤ ، وفي واقع الأمرأيضا فإن احتمال ظهورستَّتَيْن

(٦.٦) هو نفس احتمال ظهور ثلاثتين (٣.٣) ، فإن هناك حالة واحدة فقط نحصل فيها على ستتين (٦.٦) ، بينما هناك طرق متعددة نحصل فيها علي المتوسط ٣ في حالة قذف الزهرات الثلاث مرتين أو أكثر . وفي اللغة العادية نقول إن الأرقام الكبيرة من المحتمل أن تتم موازنتها بالأرقام الصغيرة خاصة عندما يكون حجم العينة كبيرًا .



الشكل رقم (١١ – ٥)

توزيع المعاينة لمتوسطات الحصول على أرقام ثلاث زهرات نرد غير عند رميها.

١١-٢ اختبار متوسط مجتمع البحث ، عندما يكون الانحراف المعياري (عـ) معلوماً :-

دعنا نري الآن كيف يمكن استخدام نظرية النهاية المركزية في الاختبارات الإحصائية . في البداية سوف نأخذ أبسط مثال ممكن للتوضيح.وحيث إن بعض الافتراضات التي نحتاج إلى طرحها في هذا النموذج غير عملية فلذا سيتم التغاضى عنها لاحقًا . إنّ كل الخطوات الخمس التي ناقشناها في الفصل العاشر سوف ندرسها بتفصيل أكبر في هذا الفصل لكي تساعدنا في التعرف أكثر علي عملية تطوير الاختبارات الإحصائية .

مسألة :- لنفترض أن باحثًا رغب في مراجعة مدي كفاءة إجراءات الحصول علي العينات المستخدمة في مسح محلي ، حيث قام بالمسح جامعو بيانات يفتقدون للخبرة اللازمة . ويشك الباحث في أن العائلات التي تنتمي إلي الشرائح ذات الدخل المتوسط والعالي قد تم تمثيلها في العينات بصورة أكبر من حقيقتها أي أنها منحت احتمالات ظهور في العينة أكبر من إحتمالات ظهور العائلات التي تنتمي إلي الشريحة ذات الدخل المتدني . إن بيانات التعداد السُكاني القائمة علي الحصر الشامل متوفرة ، وتوضح لنا أن متوسط دخل الأسرة في هذا المجتمع يبلغ ١١٥٠٠ دولار وأن الانحراف المعياري يبلغ ١٥٠٠ دولار. نفترض أن عينة صغيرة قد تم اختيارها من مجتمع بحث معين اختياراً عشوائيًا يحتوي علي ١٠٠ أسرة . فإذا كان متوسط دخل الأسرة لهذه العينة ١١٩٠٠ دولار فهل للباحث أن يشك في أن تكون هذه العينة متحيزة ؟

عمل الافتراضات :

للاستفادة من نظرية النهاية المركزية لابد من عمل بعض الافتراضات . وكما أوضحنا من قبل ، يجب أن يكون هناك دائمًا إفتراض عن منهج اختبار العينات . في هذه الحالة سوف نفترض أن العينة عشوائية . وفي الواقع أن هذا الافتراض هو الذي نود أن نختبره حيث إننا نشك في قدرة جامعي البيانات علي منح كل العائلات نفس فرص الاختيار في العينة . من المفترض أننا مستعدون لقبول بعض الإفتراضات عن مجتمع البحث . وعلى سبيل المثال نفترض أن بيانات التعداد صحيحة .وإذا لم يكن في وسعنا قبول أرقام التعداد ،فإن هناك افتراضين علي الأقل يحوم حولهما الشك ، وبالتالي فإن تفسير النتائج سيكون بالغ الصعوبة .إن عشوائيه المعاينة ستكوّن فرضنا ، وباقي الافتراضات عن مجتمع البحث ستكوِّن النموذج .وهناك حاجة لمجتمع ذي توزيع طبيعي ، إذا كان حجم العينة (ن) ليس كبيرًا . هنا يبرز السؤال إلي أي حدٍ يجب أن يكون حجم العينة كبيرًا قبل أن نقلل من الاعتماد علي افتراض التوزيع الطبيعي ، ونستفيد من نظرية النهاية المركزية . لاتوجد لدينا إجابة بسيطة على هذا السؤال ، حيث أن الإجابة على هذا السؤال تعتمد ، من بين عوامل أخري ، علي (أ) مقدار دقة تقدير احتمالية الخطأ من النوع الأول الذي نرغب فيه (ب) مامدي تطابق توزيع مجتمع البحث مع التوزيع الطبيعي . نحن وإن كنا نحذر من الأحكام المبنية فقط علي التقريرية البسيطة ، إلا أننا يمكن أن نقترح أنه كلما كان عدد أفراد العينة أكبر من أو يساوي ١٠٠ [ن ≤ ١٠٠] فإننا نستطيع أن نقلل عمليًا من الاعتماد علي افتراض التوزيع الطبيعي . فإذا كان عدد أفراد العينة (ن) أكبر من أو يساوي ٥٠ [ن ≤ ٥٠] ، وإذا كان هناك دليل تجريبي على أن الانحراف عن التوزيع الطبيعي ليس كبيرا ، فإننا نستطيع

أن نطبق الاختبارات التي نناقشها في هذا الجزء بدرجة ما من الثقة . وإذا كان عدد أفراد العينة (ن) أصغر من أو يساوي ٣٠ [ن ≥ ٣٠] يجب الا نستخدم هذه الاختبارات إلا إذا كانت درجة تطابق (تمثيل) التوزيع الطبيعي معروفة علي أنها جيدة . وكلما كان حجم العينات المستخدمة صغيراً جداً ، فإننا نعتقد دائماً مثل هذه المعلومات ، حيث لايوجد عدد كاف من أفراد العينة لتوضيح شكل توزيع مجتمع البحث ، وبالتالي كان من الأفضل تطبيق أنواع أخري من الاختبارات في حالة العينات ذات الحجم الصغير . دعنا نفترض أنه في إمكاننا أن نستخدم في هذه المسألة نظرية النهاية المركزية استخداماً صحيحاً . كما نعلم فإن توزيعات الدخل عادة ماتكون ملتوية (Skewed) ومن ناحية أخري فإن لدينا عينة كبيرة إلي حد معقول . بالاضافة إلى الأفتراضات السابقة فإننا لو أردنا تطبيق نظرية النهاية المركزية فإن من الضروري قبول أرقام التعداد بالنسبة للانحراف المعياري ، والمتوسط الحسابي لمجتمع البحث ، وأن نفترض المقاييس الفاصلة (Interval Scales)، وهكذا سيكون لدينا الافتراضات التالية :-

١ - مستوي القياس : مقاييس فاصلة

النموذج : توزيع طبيعي (يمكن التغاضى عن افتراض التوزيع الطبيعي)

المتوسط الحسابي :	١١٥٠٠دولار
الانحراف المعياري :	١٥٠٠ دولار
الفرض الصفرى (H_0)	معاينة عشوائية

٢- الحصول علي توزيع المعاينة :

Obtaining The Sampling Distribution :

من حسن الحظ أن مهمة الحصول علي توزيع المعاينة قد أنجزت بالنسبة لنا ، حيث إننا نعلم أن توزيع المعاينة لمتوسطات العينة طبيعي (أو تقريبًا طبيعي) وبذلك فإننا نستطيع أن نصل مباشرة إلي جدول التوزيع الطبيعي، ، ومنذ الآن فإن توزيعات المعاينة سوف تكون موجودة دائمًا في شكل جداول في الملحق رقم (٢) . ولكن من المهم أن ندرك أنه قد تم حساب هذه الجداول باستخدام نظرية الاحتمالات .إن من السهل جدًا أن يفقد الشخص اتجاهه في خضم تفاصيل الحسابات للدرجة التي ينسي فيها حقيقة أنه كلما استخدمنا جداول الاختبارات الإحصائية فإننا في الواقع نستخدم توزيع المعاينة .

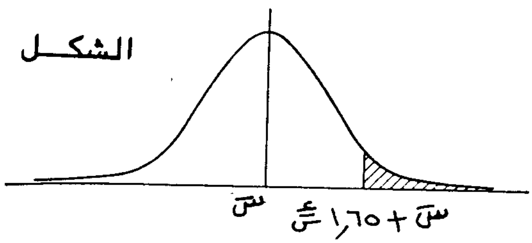
٣- اختيار مستوي الدلالة والمنطقة الحرجة :

إن إختيار مستوي الدلالة الصحيح يعتمد بالطبع علي التكاليف النسبية المرتبطة بالوقوع في الأخطاء من النوع الأول أو من النوع الثاني · إذا فشل الباحث في رفض الفرض الصِّفرى الذي مفاده أن المعاينة عشوائية بينما تكون العينة في الواقع متحيزة ، فإنه يخاطر بنشر نتائج مضللة عن العينة · ومن ناحية أخري فأنه إذا رفض الفرض في الوقت الذي يكون فيه الفرض صحيحًا ، فإنه ربما اضطر إلى إعادة المسح مع مايتضمه ذلك من تكاليف كبيرة · إن الموقف المثالي هو أن يتخذ الباحث قرارًا عقليًا مبنيًا علي أساس مقارنة التكلفة النسبية لكل نموذج من الخطأين المذكورين ·

وفي الواقع فإنّ هذه الخطوة ليست من السهولة بمكان وتحتاج إلي كثير من الجهد والوقت . دعنا نفترض أولاً: أننا قرّرناأن يكون مستوى الدلالة

٥٠و . ثانيـــاً : يــجب أن نــــقـرر اســـتخدام اخـــتبار « ذو الطـــــرف الـــواحـــد » (One Tailed Test) حيث إنَّ اتجاه التحيز قد تم التنبؤ به . إذا اتضح أن متوسط دخل العينة كان أقل من ١١٥٠٠ دولار ، فإنّ الباحث لن يستطيع أن يشك في أن جامعي البيانات قد اختاروا في العينة عدداً أكبر من أفراد فئات الدخل العالي والمتوسط (٥) . وباختيار مستوى الدلالة عند ٥٠و ، والاختبار ذو الطرف الواحد ،فإن المنطقة الحرجة يتم تحديدها من جدول التوزيع الطبيعي ، حيث إن ٥٪ فقط من منطقة المنحني الطبيعي تقع الي يمين الإحداثي السيني عند النقطة التي تبعد ١٫٦٥ انحرافا معيارياً عن المتوسط الحسابي . فإذا علمنا أن النتيجة تبعد ١٫٦٥ انحرافاً معيارياً عن المتوسط الحسابي ، فإننا سنرفض الفرض الصفرى (أنظر الشكل رقم (١١-٦) .

الشكـــل (١١ - ٦)



توزيع المعاينة الطبيعي، والمنطقة المظللة تمثل منطقة حرجة ذات طرف واحد عند مستوى دلالة يساوي ٥ و .

(٥) في هذه المسألة وعلي الرغم من أنّنا نجد البيانات عن العينة معطاه فعلاً ونعرف إتجاه النتيجة فإنّنا يجب أن نتخيل مع ذلك ، أنه قد تم اتخاذ القرار قبل معرفة النتائج .

٤- حساب الإختبار الإحصائي : -

(Computing The Test Statistics)

إنّنا نعرف أن لو كل الفروض كانت صحيحة فإن توزيع المعاينة إلى كل قيم $\bar{س}$ سيكون توزيعاً طبيعياً ($\bar{س}$ ، $\frac{ع^{٢}}{ن}$)

$Nor\ (\mu\ \frac{\sigma^2}{N})$. في إطار مثالنا يصبح :-

$\bar{س}$ = ١١٥٠٠ دولار

$ع_{\bar{س}} = \frac{ع}{\sqrt{ن}} = \frac{١٥٠٠}{\sqrt{١٠٠}} =$ ١٥٠دولار .

ولكي نستخدم جدول التوزيع الطبيعي فإن من الضروري تحويل القيم

إلي درجات معيارية أو بمعني آخر ، الحصول علي الإحصاءات (د) (Z) التي لها توزيع طبيعي ذومتوسطحسابي يساوي صفراً وانحراف معياري يساوي واحداً صحيحاً : طبيعي (صفر ،١) Nor (0, 1) . لقد استخدمنا من قبل المعادلة التالية :-

$$د = \frac{س - \bar{س}}{ع}$$

$$Z = \frac{X - \bar{X}}{S}$$

إن هذه المعادلة تصلح لعينة ذات توزيع طبيعي وسطه الحسابي $\bar{س}$ ، وتباينه ($ع^{٢}$) : طبيعي ($\bar{س}$ ، $ع^{٢}$) ، ولكن لاتصلح لتوزيع المعاينة . دعنا نتذكر مرة أخري كل خطوة من الخطوات التي استخدمنا ها .لقد افترضنا مجموعة من الإفتراضات لكي نحصل على توزيع معاينة . إن هذا التوزيع الأخير يخبرناكيفية إحتمال أن يكون وسط حسابي معين $\bar{س}$ ، إذا كانت الفروض صحيحة فعلاً . لقد حصل الباحث علي قيمة الوسط

الحسابي $\bar{س}$ منفرداً من عينته ثم بعد ذلك إستخدم نظرية توزيع المعاينة لمساعدته في تقييم احتمال الحصول علي نتيجة غير عادية ، أو غير عادية بدرجة كبيرة مقارنة مع قيمة $\bar{س}$ التي حصل عليها .

إن الباحث يتعامل مع توزيع المعاينة عندما يستخدم جدول التوزيع الطبيعي . إن كل حالة من هذا التوزيع هي ($\bar{س}$) المتوسط الحسابي للعينة ، وأن المتوسط الحسابي لهذا التوزيع هو $\bar{\bar{س}}$ والانحراف المعياري هو $\frac{ع}{\sqrt{ن}}$.

إذن إن قيمة $\bar{س}$ تحل محل س ، وقيمة $\bar{\bar{س}}$ تحل محل $\bar{س}$ ، و $\frac{ع}{\sqrt{ن}}$ تحل محل (ع) في المعادلة السابقة للدرجة المعيارية (د) .

$$\text{إذن} \qquad د = \left(\frac{\bar{س} - \bar{\bar{س}}}{\frac{ع}{\sqrt{ن}}} \right)$$

$$Z = \frac{X - \mu}{\frac{\sigma}{\sqrt{N}}}$$

$$د = \frac{١١٩٠٠ - ١١٥٠٠}{١٥٠} = ٢٫٦٧$$

وبمعنى آخر فإن متوسط العينة يزيد بمقدار ٢,٦٧ خطأً معيارياً عن متوسط مجتمع البحث.

٥ – اتخاذ القرار : –

وبما أن ($\bar{س}$) المتوسط الحسابي للعينة ينحرف عن متوسط مجتمع البحث المفترض $\bar{\bar{س}}$ في الاتجاه المفترض بأكثر من ١,٦٥ انحرافًا معياريًا، فإنه يجب رفض الفرض عند مستوي الدلالة ٠,٠٥ وفي الواقع أنّنا لو حسبنا

الدرجة المعيارية (د) بالضبط فإنّنا نستطيع أن نقول أكثر من ذلك .إنّ احتماال الحصول على (د) بهذا الحجم أو أكبر هو ٠٠٣٨, ، باستخدام اختبارٍ ذي طرف واحد . وعمليًا فإن من الأفضل عادة حساب مستوي الدلالة بالضبط كلما كان ذلك ممكنًا . وإذا قمنا بهذا العمل ، فإننا نوضح أن النتيجة تقع تحت منطقة حرجة أكثر صغرًا من التي أوضحناها من قبل. وحيث إن القارئ ربما يفضل أن يستخدم مستوي دلالة مختلف عن الذي استخدمه المؤلف ، فإنه من المفيد تقديم احتمالات دقيقة أو أكثر قربًا من الدقة حتي يستطيع القارئ أن يستخلص لنفسه النتائج التي يراها حول ماإذا كان سيقبل أو يرفض النتائج التي تم التوصل إليها . وفي هذا المثال فإن الباحث سوف يرفض الفرض الصفرى بأن المعاينة عشوائية .

ثم عليه بعد ذلك أن يقرر ما إذا كان سيقوم باختيار عينة أخري أم لا .

١١ - ٣ توزيع ستيودنت التائى : -

Student' s ((t)) Distribution

في معظم الحالات يكون من غير الواقعي أن نعامل الانحراف المعياري لمجتمع البحث وكأنه معلوم . في العادة أن الباحث يبذل جهدًا صعبًا لتأكيد العشوائية ، حيث أنه يكون مهتمًا باختبار الافتراضات الخاصة بمجتمع البحث الذي تمت دراسته . وفي الاختبارات من النوع الذي نناقشه في هذا الفصل فإن من المحتمل أن يختبر فرضًا يتصل بمتوسط مجتمع البحث. ولكن إذا كانت هذه هي الحالة ، هل سنكون في موقف نعرف فيه قيمة الانحراف المعياري (عـ) (σ) ؟ عمليًا يستحيل أن نعرف ذلك ، لانه لو كانت لدينا أي معلومات عن الانحراف المعياري لمجتمع البحث فسنكون

بلاشك في موقف نعرف فيه قيمة المتوسط الحسابي لمجتمع البحث ($\bar{س}$) ، أو الانحراف المعياري لمجتمع البحث ، إلا إذا كان هناك شخص ما مثل الإحصائي الفني الذي نتعامل معه ، يتعمدُ أن يحجب عنا المعلومات . ولكن في العادة لن نعرف قيمة المتوسط الحسابي لمجتمع البحث ($\bar{س}$)أو الانحراف المعياري لمجتمع البحث (عـ) . فما الذي نستطيع أن نفعله في مثل هذا الموقف ؟ بما أنّ نظرية النهاية المركزية تتضمن الانحراف المعياري، فلا يمكن تجاهل قيمته كلية .إن واحداً من الحلول هو استبدال الانحراف المعياري لمجتمع البحث (عـ) بالانحراف المعياري للعينة (ع) . وفي واقع الأمر فإن هذا الأسلوب كان متبعًا دائمًا قبل تطور الإحصاء الحديث . وفي معادلة الدرجة المعيارية (د) قد تم استبدال $\frac{عـ}{\sqrt{ن}}$ بالمعادلة $\frac{ع}{\sqrt{ن}}$

وبما أنه من الممكن حساب (ع) مباشرة من بيانات العينة فلذا لانجد أي مجهول في المعادلة . وكما يتضح فإن هذه الخطوات تؤدي إلي نتائج جيدة نسبيًا عندما يكون عدد الأفراد (ن) كبيراً . وكما سنري بعد ذلك فإن الاحتمالات التي نحصل عليها بهذه الطريقة تكون خادعة كلما كانت (ن) صغيرة نسبيًا .دعنا نري لماذا يحدث هذا .

يمكننا أن نبني إختبارًا إحصائيًا بديلاً

$$ت = \frac{\bar{س} - \bar{س}}{\frac{ع}{\sqrt{ن - ١}}}$$

$$t= \frac{\bar{X} - \mu}{\frac{S}{\sqrt{N - 1}}}$$

لقد قام « جوست » بابتكار هذه المعادلة وسماها باسم «ستيودنت » ولها توزيع معاينة باسم توزيع «ستيودنت التائي» Student's t - Distribution

فإذا ما قارنا بين (ت) و (د) يتضح لنا الاتي :-

بينما يتشابه البسط في المعادلتين فإن المقامات تختلف من ناحيتين أولاً :- هناك (ن - ١) تحت الجذر ، ثانيًا : تم استبدال الانحراف المعياري لمجتمع البحث في معادلةالدرجة المعيارية (د) بالانحراف المعياري للعينه في معادلة القيمة التائية (ت) .ولكي نفهم هذه التعديلات دعنا نقوم بفحص كلّ منهما بالتتابع . ولكي نقوم بهذا العمل يجب أن ندخل عدة أفكار جديدة .

إن الانحراف المعياري للعينة يمكن أن يستخدم كتقدير للانحراف المعياري لمجتمع البحث . وعلي الرغم من أن مشكلة التقديرات سوف تدرس في الفصل القادم فإنه يكفي أن نذكر هنا أننا غالبًا مانريد للتقدير أن يمتلك بعض الخصائص . إن واحداً من خصائص التقدير الجيّد ألاّ يكون متحيزًا . وعلي عكس مايمكن توقعه فإننا نجد أنّ (ع) ليست تقديرًا غير متحيز تمامًا للانحراف المعياري لمجتمع البحث وهناك قيمة أخري يمكن الإشارة إليها بالرمز ($\hat{ع}$)($\hat{\sigma}$) والتي نحصل عليها بالمعادلة التالية :-

$$\hat{ع} = \sqrt{\frac{\sum_{و=١}^{و=ن} (س_و - \bar{س})^{٢}}{(ن - ١)}}$$

$$\hat{\sigma} = \sqrt{\frac{\sum_{i=1}^{N} (x_i - \bar{X})^2}{(N-1)}}$$

ويمكن أن نوضح رياضيًا أنّ هذه القيمة $\hat{ع}$ ($\hat{\sigma}$) تقدير غير

متحيز لقيمة عـ (σ) (الانحراف المعياري لمجتمع البحث) [٦] . إن الفرق الوحيد بيني (عـ) و (ع) هو في قيمة المقام (ن-١) . إذن فبالرغم من أنك قد تعلمت كيف تحسب (ع) ، تجد نفسك الآن مواجه بحقيقة أن معادلة أخري يجب استخدامها في تقدير (عـ). وفي المسألة الحالية فإننا نحتاج لتقدير $\frac{\hat{ع}}{\sqrt{ن}}$ (وليس لتقديرعـ) ، وحيث إن الأولي هي التي تظهر في مقام معادلة الدرجة المعيارية ، وعلى الرغم من أن من الصحيح أن $\frac{\hat{ع}}{\sqrt{ن}}$ هو أفضل تقدير لقيمة $\frac{\hat{ع}}{\sqrt{ن}}$ فإن من الممكن أن نتفادي كلية حساب (عـ) من (عـ) ، إذا كنا قد حصلنا فعلاً على (ع) . لاحظ الآتي

$$\frac{\hat{ع}}{\sqrt{ن}} = \frac{\sqrt{\left[\sum_{و=١}^{و=ن} (س_و - \bar{س})^٢\right] / (ن-١)}}{\sqrt{ن}}$$

$$\frac{\hat{\sigma}}{\sqrt{N}} = \frac{\sqrt{\left[\sum_{i=1}^{i=N} (x_i - \bar{x})^2\right] / (N-1)}}{\sqrt{N}}$$

ويجب أن نتذكر أنّ $\sqrt{أ} \div \sqrt{ب} = \sqrt{\frac{أ}{ب}}$

---

(٦) لكي نكون دقيقين فإنّ $\hat{ع}$ ($\hat{\sigma}$) ليستْ تقديرًا غير متحيز لقيمةعـ(σ) ، ولكن قيمة $\hat{ع}^٢$ هي تقدير غير متحيز لقيمة $عـ^٢$. ولكن لايجب أن يشغلنا هذا التمييز الدقيق . في هذا الكتاب سوف نستخدم دائما العلامة(٨) علي حرف يوناني لكي تشير إلي تقدير المعلّم (Parameter) . بعض الكتب تعرف الانحراف المعياري للعينة (ع) في المقام بقيمة (ن-١) ، ولكننا سوف نجعل كل معادلةمن المعادلتين منفصلة عن الأخري .

وعليه ، $$\frac{\hat{\text{ع}}}{\sqrt{\text{ن}}} = \sqrt{\frac{\sum_{\text{و}=1}^{\text{و}=\text{ن}} (\text{س}_{\text{و}} - \bar{\text{س}})^2}{\text{ن}(\text{ن}-1)}}$$

$$\frac{\hat{\text{ع}}}{\sqrt{\text{ن}}} = \frac{\sqrt{\sum_{\text{و}=1}^{\text{و}=\text{ن}} (\text{س}_{\text{و}} - \bar{\text{س}})^2 / \text{ن}}}{\sqrt{(\text{ن}-1)}} = \frac{\text{ع}}{\sqrt{\text{ن}-1}}$$

وهكذا فإننا نستطيع أن نأخذ تقديراً متحيزاً الي حدٍ ما لقيمة (عـ) ونقسمها على قيمة أقل قليلاً من $\sqrt{\text{ن}}$ ونحصل علي قيمة $\frac{\text{ع}}{\sqrt{\text{ن}-1}}$ لتقدير غير متحيز لقيمة $\frac{\text{عـ}}{\sqrt{\text{ن}}}$.

ولهذا السبب فإن قيمة (ن-١) تظهر في مقام معادلة التوزيع التائي .[٧]

عند استبدال (د) بالقيمة التائية (ت) نجد أن التعديل الذي تم إدخاله باستخدام (ن - ١) يكون سطحيًا (نسبيًا) ، ولكن وضع (ع) في محل (عـ) ربما يكون ذا دلالة أكبر إذا كانت قيمة (ن) صغيرة ٠ وحيث إن (ع) تختلف من عينة إلي أخري فإن مقام التعبير التائي يختلف كما هو الحال في البسط ٠ ولقيمة معينة للمتوسط الحسابي ($\bar{\text{س}}$) ، إذا كان الانحراف المعياري لعينة ما صغيراً جداً فإن القيمة التائية ستكون كبيرة جداً . وإذا كان الانحراف المعياري (ع) كبيرًا ، فإن القيمة التائية ستكون صغيرة نسبيًا.

---

(٧) بعض الكتب توصي باستخدام (ن -١) للعينات الصغيرة و(ن) للعينات الكبيرة ٠ ويبدو أن هذا الإجراء يضيف تعقيداً لالزوم له ٠ في حالة العينات الكبيرة بالطبع لايوجد فرق كبير في أي من التعبيرين نستخدم ٠

وعليه ستكون هناك إختلافات كبيرة وسط الدرجات التائية أكثر من اختلافات الدرجات المعيارية المقابلة . وهذا يعني أن التوزيع التائي سيكون مسطحًا أكثر من التوزيع الطبيعي . كما سيكون للتوزيع التائي نهايات كبيرة (Larger Tails) . وإلي أي مدي سيكون التوزيع التائي مسطحًا يعتمد على حجم العينة .إذا كان حجم العينة(ن) صغيرًا جدًا فإن التوزيع التائي سيكون مسطحًا جدًا مقارنة مع التوزيع الطبيعي . وبمعنى آخر ، يصبح من الضروري الابتعاد أعدادًا كثيرة من الانحرافات المعيارية عن المتوسط الحسابي لنضم ٩٥٪ من المفردات . وكلما أصبح حجم العينة (ن) كبيرًا فإن التوزيع التائي يقترب أكثر فأكثر من التوزيع الطبيعي ولكن مع ذلك يكون إلي حدما أكثر تسطحًا من التوزيع الطبيعي . وهكذا فإن هناك توزيعًا تائيًا خاصًا لكل حجم عينة .وحقيقة أن التوزيعات التائية تقترب من التوزيع الطبيعي تبدو معقولة عندما ندرك أنه كلما أصبح حجم العينة (ن) كبيرًا ، فإن الانحراف المعياري للعينة (ع) يصبح تقديرًا دقيقًا جدًا للانحراف المعياري لمجتمع البحث (عـ) ، ويصبح الفرق صغيرًا جدًا إذا ما استخدمنا «ع» أو «عـ» في المقام .

للاستفادة من التوزيع التائي يجب إفتراض أن مجتمع البحث موزع توزيعًا طبيعيًا ، خاصة إذا كان حجم العينة (ن) صغيرًا نسبيًا . إن حسابات توزيع المعاينة التائي (t) تحتاج إلي أن يكون البسط ($\bar{س} - \bar{\bar{س}}$) موزعًا توزيعًا طبيعيًا ، وإن قيم البسط تكون مستقلة عن قيم المقام $\frac{ع}{\sqrt{ن - ١}}$ وعادة نحن لانتوقع استقلالاً بين البسط والمقام لأنه من المدهش أن يكون المتوسط الحسابي ($\bar{س}$) والانحراف المعياري (ع) للعينة مستقلان إستقلالاً

إحصائيًا كل منهما عن الأخر . وإذا علمنا المتوسط الحسابي للعينة ($\bar{س}$) ، فإننا نتوقع تحسن قدرتنا علي التنبؤ بالانحراف المعياري (ع) لنفس العينة . ويحدث أيضًا أنّه بالنسبة لمجتمع التوزيع الطبيعي والمعاينة العشوائية فإن متوسط العينة والانحراف المعياري سيكونان مستقلين إحصائيًا . وبما أن هذه الخاصية لاتنطبق علي كل توزيعات مجتمع البحث ، وبما أن ($\bar{س}$ − $\bar{\bar{س}}$) ، ($\bar{X}$ - μ) لاتتوزع عادة توزيعًا طبيعيًا مالم يكن حجم مجتمع البحث كبيرًا ، لهذا يجب أن نفترض مجتمعًا ذا توزيع طبيعي عند استخدام الاختبار التائي (t - Test) .

مســـألة :

دعنا نفترض أنّنا نقوم بتقييم برامج خمسة وعشرين مركزًا للخدمة الإجتماعية في ولاية نيويورك الأمريكية ، تم إختيارها عشوائيًا من بين كل مراكز الخدمة الاجتماعية في الولاية . وأن كل مركز يحتفظ بسجل يرصد فيه النسبة المئوية للحالات الناجحة حسب أسس تقييم تم الاتفاق عليها . لقد وُضِعَ معيارٌ يقيس النجاح الذي يكون عند مستوي ٦٠٪ لكل المراكز في عينة البحث ، ووجدنا أن متوسط النسبة المئوية يكون ٥٢٪ ، وأن الانحراف المعياري يكون ١٢٪ . هل لديك أي سبب يجعلك تشك في أنه بالنسبة لمجتمع المراكز ككل أن مستوي الأداء أقل من المعيار المتوقع ؟

١- وضع الافتراضات :-

إن الافتراضات الضرورية يمكن أن تكون علي النحو التالي :-

النمــوذج : معاينة عشوائية ، وتوزيع مجتمع البحث عشوائيًا .

مستوي القياس : مقاييس الفترة .

الفرض : متوسط مجتمع البحث = س = ٦٠٪ .

لاحظ عدم وجود افتراض بخصوص الانحراف المعياري لمجتمع البحث (عـ) ، حيث إن الانحراف المعياري للعينة (ع) تم الحصول عليه تجريبياً ، ويمكن استخدامه مباشرة في الإختبار التائي . ونحتاج هنا إلي بعض الملاحظات عن مستوي القياس . بما أن كل عميل لمركز يمثل للمركز نجاحاً أو فشلا ، وحيث إن الأرقام التي حصلنا عليها لكل مركز هى نسب مئوية للنجاح ، ربما يعتقد البعض أننا نتعامل مع مقياس إسمي ثنائي الفئات ، (Dichotomized Nominal Scale) بدلاً من مقياس الفترة. . بالطبع لو كانت وحدات التحليل عملاء وليست مراكز لكان هذا الأمر صحيحاً . ولكن تذكر أن وحدات االتحليل هنا مراكز . لقد حصلنا على درجة « نسبة مئوية للنجاح » لكل مركز، وهذه الدرجة تمثل مقياس فترة صحيح . وعلي سبيل المثال ، فإن الفرق بين ٣٠٪ و ٤٠٪ هو نفس الفرق بين ٦٠٪ و ٧٠٪ . إن كلا الفرقين يمكن ترجمتهما إلي نفس العدد الحقيقي للعملاء .

٢- الحصول على توزيع المعاينة :

الجدول رقم (د) في الملحق رقم (٢) يوضح توزيعات المعاينة التائية (t) ، وبما أن هذه التوزيعات تختلف حسب حجم كل عينة ، فقد تم تركيز الجدول بحيث تعطي فقط أطراف كل توزيع . وعند استخدام الجدول من الضروري أولاً : تحديد الحجم المناسب لكل عينة وذلك بالنظر إلي العمود الواقع علي اليسار . إن أحجام هذه العينات موجودة على شكل درجات حرية (Degrees Of Freedom) ، والتي في مثل هذا النوع من المسألة تساوي دائماً

(ن – ١)[٨] ، ثم بعد ذلك حدد مستوي الدلالة المناسب وذلك بأن تقرأ من أعلى . إن الأرقام الموجودة في داخل الجدول تشير إلي القيم التائية المطلوبة للحصول علي مستوي الدلالة المحدد .

٣- اختيار مستوي الدلالة والمنطقة الحرجة :

دعنا نستخدم مستوي الدلالة عند القيمه ٠٥,٠ ، ونستخدم اختبار الطرف الواحد (One - Tail Test) . ففي الجدول رقم (د) (الملحق رقم (٢)]، نري أن من الأربع والعشرين درجة من درجات الحرية ، فإننا نحتاج الي قيمة تائية تساوي ٢,٠٦٤ أو أكبر ، بهدف الحصول علي دلالة عند مستوي ٠٥,٠ لاختبار ذي الطرفين . أما الاختبار ذي الطرف الواحد وبمستوي دلالة ٠٥,٠ ، فإننا نحتاج لقيمة تائية تساوي ١,٧١١ أو أكبر . في حالة الاختبارات ذات الطرف الواحد فإننا بكل بساطة قد قمنا بتصنيف مستويات الدلالة الضرورية لاختبارات الطرفين . وهذا راجع إلي أنّنا نبتعد نفس عدد الانحرافات المعيارية عن المتوسط الحسابي بهدف الحصول علي منطقة حرجه ذات طرف واحد قيمتها ٠٥,٠ كما لو كنا سنحصل علي منطقة حرجة ذات طرفين تساوي ١٠و.

٤- حساب إحصاءات الاختبار :

Computing Test Statistics.

علي الرغم من أن توزيع المعاينة للمتوسط الحسابي (سَ) هو توزيع طبيعي حقًا طبيعي (سَ ، ع٢/ن) ، وبالتالي فإن توزيع الدرجات المعيارية (د) هو الأخر توزيع طبيعي بمتوسط حسابي «صفر » وانحراف معياري

(٨) لمناقشة درجات الحرية انظر القسم (١٢ – ١).

واحد صحيح : طبيعي (صفر،١) ، فإن هذه المعلومة لاتفيدنا فائدة حقيقية ، إذ أن قيمة الانحراف المعياري (ع) مجهولة .وبدلاً من ذلك فإننا نحسب القيمة التائية بالمعادلة التالية :

$$ت = \frac{\bar{س} - س}{\frac{ع}{\sqrt{ن - ١}}} = \frac{٥٢ - ٦٠}{\frac{١٢}{\sqrt{٢٤}}} = -٣٫٢٧$$

$$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{N - 1}}} = \frac{52 - 60}{\frac{12}{\sqrt{24}}} = -3.27$$

٥- القرار:

لقد تحدد أن أي قيمة تائية أكبر من أو تساوي ١,٧١١ [ت ≤ ١,٧١١] ستقع داخل المنطقة الحرجة . ولهذا سوف نرفض الفرض الصفرى الذي مفاده أن متوسط مجتمع البحث (س) يساوي.٦٪ ونستنتج - مع بعض مخاطر الخطأ - أن مستوي الأداء الفعلي للمراكز أقل من المستوي المتوقع . وبالاطلاع علي الصف المواجه للدرجة ٢٤ من درجات الحرية في الجدول رقم (د) في الملحق رقم (٢) ، نري أن مستوي الدلالة لاختبار ذي الطرف الواحد المماثل للقيمة التائية ٣,٧٢ يقع مابين ٠,٠٠٥ و ٠,٠٠٠٥(٩).

(٩) علي الرغم من أن الاحتمالات الدقيقة لايمكن الحصول عليها من الجدول إلا أن الاستكمال (In-terpolation) يكون دائمًا ممكنًا . عادة يكفي أن نشير إلي أن الاحتمال (ح) يقع بين قيمتين، علي سبيل المثال ٠,٠٠٥ ≤ ح ≤ ٠,٠٠٠٥

عند هذه النقطة يمكن أن نلاحظ عدداً من الحقائق حول التوزيع التائي . إذا نظرنا للعمود الذي يوازي (ح) = ٠٥, ٠ : P =.05 } لاختبار ذي الطرفين ، فسوف نلاحظ أنه كلما كبر حجم العينة ، فإن القيمة التائية تصبح أصغر وتقترب بسرعة من ٩٦ر١ وهي القيمة الضرورية للدلالة إذا استعملنا جدول التوزيع الطبيعي .إن هذه القيمة يجب أن تعطي فكرة جيدة عن درجة التقريب إلي المنحني الطبيعي لحجم أي عينة . بالنسبة لقيم (ن-١) التي تكون أكثر من ٣٠، فإن عملية الاستكمال (Interpolation) تصبح ضرورية في العادة . أما بالنسبة للقيم التي تكون أكثر من ١٢٠ فيجب استخدام جدول التوزيع الطبيعي ، حيث إن القيم التائية غير متوفرة في هذه الحالة . إن بعض الكتب تذكر تعسفيًا أنه من الضروري استخدام جدول التوزيع التائي إذا كانت قيمة (ن) أصغر من أو تساوي ثلاثين [ن ≤ ٣٠] . وبالرغم من أن هذا الحكم التقريري يعطينا نتائج معقولة ، لكن الموقف الذي نأخذ به هنا هو أنه من الأفضل دائمًا استخدام الجدول التائي عندما يكون الانحراف المعياري (عـ) مجهولاً ، وعندما يكون من الممكن افتراض مجتمع بحث ذي توزيع طبيعي .وحيث أن الجدول التائي لم يعد استعماله صعبًا ، فإنه يبدو من الأفضل استخدام القيم الدقيقة بدلاً من التقريب العادي . يجب أن نؤكد هنا أنه لاتوجد نظرية فريدة تطبق علي كل العينات الصغيرة وأخري مختلفة تمامًا تطبق علي كل العينات الكبيرة كما توصي بذلك بعض الكتب .

وكما هو واضح من جدول التوزيع التائي ، فإن التوزيع الطبيعي والتوزيع التائي يختلفان إختلافًا كبيراً فقط في حالة ماإذا كان حجم العينة صغيرًا نسبيا .

كما أننا يجب أن نفترض توزيعًا طبيعيًا كلما استخدمنا التوزيع التائي ، مالم يكن حجم العينة كبيرًا جدًا . وفي هذه الحالة يمكن تقريب القيمة التائية بقيم الدرجات المعيارية (د) . وهكذا فإن القيمة العملية للاختبار التائي نجدها في الحالات التي تكون فيها العينات صغيرة ، وعندما يمكن افتراض مجتمع ذي توزيع طبيعي . ومن سوء الحظ ، فإنه عندما تكون العينات صغيرة الحجم نصبح عادة في شك كبير حول حقيقة التوزيع الطبيعي لمجتمع البحث . على سبيل المثال إذا كان الباحث يقوم بعمل دراسة استطلاعية مستخدمًا عينة من ١٧حالة . هل يكون الباحث في موقف يقبل فيه افتراض مبدأ التوزيع الطبيعي ؟

في الغالب لايستطيع الباحث أن يقبل الافتراض . وكما سنري في الجزء الثاني من هذا الكتاب فإن هناك اختبارات أخري يمكن أن تستخدم كبدائل للاختبار التائي ولاتتضمن افتراض التوزيع الطبيعي .

١١-٤ اختبارات تتضمن التناسب :-

Tests Involving Proportions.

حتي الآن تعاملنا ، في هذا الفصل مع الأمثلة التي تتضمن مقاييس الفترة فقط · وكذلك كان يجب افتراض التوزيع الطبيعي لمجتمع البحث في حالة العينات الصغيرة · إن في هذا الجزء من الفصل سوف نري كيف يمكن استخدام نظرية النهاية المركزية لتشمل اختبارات تتضمن التناسب عندما تكون قيمة (ن) كبيرة الي حدما · وفي الواقع فإننا سنعامل التناسب كحالات خاصة من المتوسطات حتي يمكن لنقاشنا السابق أن يظل قائمًا . فلنفترض أن لدينا مقياسًا اسميًا ثنائي الفئات ، وأننا نود أن نختبر فرضًا خاصًا بنسبة الذكور في مجتمع ما . دعنا نحدد اعتباطًا الرقم واحد للذكور

والرقم (صفر) للإناث ، وأن نعامل الدرجات كمقاييس فترة . وبالرغم من عدم وجود وحدة محددة بوضوح، ما لم تكن صفة الذكورة التي تكون موجودة أو غير موجودة ، فإننا يمكن أن نعامل هذه الأرقام الاعتباطية كمقاييس فترة لأنه يوجد فقط وحدتين من الأرقام . أما إذا أضفنا فئة ثالثة لفئات المتغير الا سمي فإن هذا لم يعد ممكنا بعد الآن - حيث يكون من الضروري أن نحدد بالضبط وضع هذه الفئة بالنسبة للفئتين السابقتين .إن مانقوله في الواقع هو أن لاضرورة للتمييز بين المقاييس الاسمية ومقاييس الرتبة ومقاييس الفترة في حالة وجود ثنائية الفئات حيث أن مسألة مقارنة المسافات بين الدرجات لاتواجهنا أبداً .

لدينا الآن مجتمع مكون بأكمله من الفئات واحد ، و صفر ، وهذا توزيع ذو منوالين حيث تتمركز كل الحالات في واحدة من النقتطين . ومن المؤكد أن التوزيع ليس توزيعاً طبيعياً . ولكن إذا كانت (ن) كبيرة فإننا نعرف أنه بغض النظر عن شكل مجتمع البحث ، فإن توزيع المعاينة لمتوسطات العينة سيكون كالاتي تقريباً : طبيعـــي (سَ ، $\frac{ع^{٢}}{ن}$) ($Nor\left(\mu, \frac{\sigma^2}{N}\right)$

وكل ماتبقي من الخطوات هو تحديد المتوسط والانحراف المعياري لهذا المجتمع المكون من الأرقام واحد ، و صفر (Ones & Zeros) .

دع «$ح_{سَ}$» (P_u) تمثل نسبة الذكور في مجتمع البحث وَك $_{سَ}$ (q_u) تمثل نسبة الإناث ، وأن الرمز (س) الموجود في أسفل رمز النسبة يشير إلى أننا نتعامل مع المجتمع بأكمله . ولكي نحصل على متوسط الآحاد

(Ones) ومتوسط الأصفار (zeros) بالنسبة لمجتمع البحث ، فإننا نقوم بجمع القيم ونقسم علي العدد الكلي للحالات . إن عدد الآحاد (Ones)

سيكون مجموع العدد الكلي للحالات مضروباً في نسبة الذكور . وبغض النظر عن عدد الأصفار فإن مساهمتهم في المجموع ستكون صفراً . وعليه فإن متوسط مجتمع البحث سيكون علي النحو التالي :-

$$\text{سَ} = \frac{\text{م ح}_{\text{س}}}{\text{م}} = \text{ح}_{\text{س}}$$

$$\mu = \frac{M P_\mu}{M} = P_\mu$$

حيث (م) تمثل حجم مجتمع البحث (مع ملاحظة الفرق بين حجم مجتمع البحث وحجم العينة (ن) . وإذن فإن متوسط عدد من الآحاد (Ones) والأصفار هو بالضبط نسبة الآحاد . وباستدلال مماثل فإن

(سَ = حَ) [$(\bar{X} = P_S)$ حيث تمثل حَ نسبة الذكور في العينة، وباستخدام المعادلة العامة للإنحراف المعياري نستطيع أن نوضح التالي :

$\text{ع} = \sqrt{\text{ح}_{\text{س}} \times \text{ك}_{\text{س}}}$ وباستخدام الرموز الخاصة بمعالم مجتمع البحث فإن معادلة الانحراف المعياري ع ، (σ) تصبح علي النحو التالي :-

$$\text{ع} = \sqrt{\frac{\sum_{\text{و}=1}^{\text{م}} (\text{س}_{\text{و}} - \text{سَ})^2}{\text{م}}} = \sqrt{\frac{\sum_{\text{و}=1}^{\text{م}} (\text{س}_{\text{و}} - \text{ح}_{\text{س}})^2}{\text{م}}}$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{M} (X_i - \mu)^2}{M}} = \sqrt{\frac{\sum_{i=1}^{M} (X_i - P_\mu)^2}{M}}$$

عندما ننظر إلي بسط الكمية تحت الجذر التربيعي فإننا نري أنه سيكون هناك نوعان فقط من القيم تمثل مربع الانحرافات عن المتوسط الحسابي

$ح_{\bar{س}}$ (P_μ) . ولكل قيمة للرقم واحد سيكون مربع الانحراف عن المتوسط الحسابي ($١ - ح_{\bar{س}}$)^٢ . ولكل قيمة للرقم صفر سيكون مربع الانحراف عن المتوسط الحسابي (صفر − $ح_{\bar{س}}$)^٢ . وبما أنه سيكون هناك م × $ح_{\bar{س}}$ من الآحاد ، م × $ك_{\bar{س}}$ من الأصفار من مجموع المربعات فإننا نحصل علي المعادلة التالية :

$$ع = \sqrt{\frac{م\,ح_{\bar{س}}(١ - ح_{\bar{س}})^٢ + م\,ك_{\bar{س}}(صفر - ح_{\bar{س}})^٢}{م}} = \sqrt{\frac{م\,ح_{\bar{س}}\,ك_{\bar{س}}^٢ + م\,ك_{\bar{س}}\,ح_{\bar{س}}^٢}{م}}$$

$$\sigma = \sqrt{\frac{M P_\mu (1 - P_\mu)^2 + M q_\mu (0 - P_\mu)^2}{M}} = \sqrt{\frac{M P_\mu q_\mu^2 + M q_\mu P_\mu^2}{M}}$$

لاحظ أنه بتحليل التعبير (م $ح_{\bar{س}}$ $ك_{\bar{س}}$) من كل حد في البسط سنحصل علي المعادلة التالية : −

$$ع = \sqrt{\frac{م\,ح_{\bar{س}}\,ك_{\bar{س}}(ك_{\bar{س}} + ح_{\bar{س}})}{م}} = \sqrt{\frac{م\,ح_{\bar{س}}\,ك_{\bar{س}}}{م}} = \sqrt{ح_{\bar{س}}\,ك_{\bar{س}}}$$

بما أن (م) تُلغي في كل المعادلتين لكل من المتوسط الحسابي ($\bar{س}$) والإنحراف المعياري (ع) ، فإن المتوسط الحسابي والانحراف المعياري لمجتمع البحث يصبحان مستقلين عن الحجم الحقيقي لمجتمع البحث كما كنا نتوقع .

إذن نستطيع أن نستخدم نظرية النهاية المركزية لكي تعطينا المعادلة التالية : -

$$ع_{\bar{س}} = ع_{ح_{\bar{س}}} = \sqrt{\frac{ع}{ن}} = \frac{\sqrt{ح_{س} ك_{س}}}{\sqrt{ن}} = \sqrt{\frac{ح_{س} ك_{س}}{ن}}$$

$$\sigma_{\bar{x}} = \sigma_{P_s} = \frac{\sigma}{\sqrt{N}} = \frac{\sqrt{P\mu\, q\mu}}{\sqrt{N}} = \sqrt{\frac{P\mu\, q\mu}{N}}$$

حيث يشير الرمز (ع $_{ح_{\bar{س}}}$) (σP_s) إلي أننا نتعامل مع الخطأ المعياري لنسب العينة . وفي مصطلحاتنا الجديدة فإن (ح $_{\bar{س}}$) تحل محل ($\bar{س}$) . كما أن(ح $_{س}$) تحل محل(س)، كما أن (ع $_{ح_{\bar{س}}}$) تحل محل (ع $_{\bar{س}}$) في معادلة الدرجة المعيارية (د) . وعليه تصبح معادلة (د) علي النحو التالي :

$$د = \frac{\bar{س} - س}{ع_{\bar{س}}} = \frac{(ح_{\bar{س}} - ح_{س})}{\sqrt{(ح_{س}\ ك_{س}) \div ن}}$$

$$Z = \frac{\bar{X} - \mu}{\sigma_{\bar{x}}} = \frac{P_s - p\mu}{\sqrt{P\mu\, q\mu / N}}$$

لاحظ أنه يبدو أن لدينا معادلة تختلف تمامًا عن المعادلة السابقة التي استخدمناها ، ولكن لاجديد في الواقع سوي تغيير في الرموز فقط ، وهذا صحيح ، حيث إننا تمكنا من أن نوضح أن النسب يمكن معادلتها وكأنها

حالات خاصة من المتوسطات . ولكن يجب أن نوضح هنا أن نظرية النهاية المركزية تحتاج إلي أن تكون (ن) كبيرة ، وذلك لكي نستفيد من تقريب التوزيع الطبيعي .وكلما كان حجم العينة (ن) صغيراً ، فإنّ نظرية ذات الحدين تكون أكثر ملاءمة .

إن هناك صلة وثيقة بين هذا الاختبار المتضمن النسب وتوزيع ذي الحدين. لقد أشرنا إلي أنه إذا كانت (ن) كبيرة ، وإذا كانت قيمة (ن × ح) أكبر من خمسة ($NP > 5$) ، حيث قيمة (ح) أصغر من قيمة (ك) ($p < q$) ، فإن توزيع ذي الحدين يمكن تقريبه بالتوزيع الطبيعي . وبالطبع فإنه في حالة التوزيع ذي الحدين فإننا نتعامل مع عدد مرات النجاحات (Number Of Successes) وليس مع التناسبات (Proportions) . إن القيمة المتوقعة لأعداد النجاحات تصبح ن ح (N P) ، والانحراف المعياري لعدد النجاحات يصبح

$$\sqrt{\text{ن ح ك}} \quad [\sqrt{NPq}]$$

ولتحويل كلّ من هذه القيم إلى قيم تناسب (Proportions) يمكننا أن نقسم كلاً منها على (ن) ، فنحصل علي قيمة التناسب (ح) للقيمة المتوقعة ، كما نحصل على المعادلة التالية :-

$$\frac{1}{\text{ن}}\sqrt{\text{ن ح ك}} = \sqrt{\frac{\text{ن ح ك}}{\text{ن}^2}} = \sqrt{\frac{\text{ح ك}}{\text{ن}}}$$

$$\frac{1}{N} \quad \sqrt{NPq} = \sqrt{\frac{N P q}{N 2}} = \sqrt{\frac{P q}{N}}$$

للحصول على الإنحراف المعياري .

وهكذا ففي حالة العينات الكبيرة سيكون بإمكاننا أن نعد مسألة خاصة بتوزيع ذي الحدين في صورة التناسب بعد تغيير رموزنا إلي ح س

وَ ك س $(P_\mu \ , \ Q_\mu)$ ، ثم نعالج المسألة وفقًا للخطوات الموضحة في هذا الفصل . علي سبيل المثال في حالة اختبارات الإشارة (Sign Test) يمكننا أن نستخدم الفرض الصفري الذي مفاده أن ح س = ٥و٠

$(P_\mu = .5)$ ، ونقارن هذه القيمة مع تناسب النجاحات التي وجدناها فعلاً في العينة .

مسألة :-

إذا كان لدي الشخص رغبة في تقييم وكالة خدمة إجتماعية معينة ، واختار من ملفاتها عينة عشوائية من ١٢٥ حالة . وقد وجدأن النسبة المئوية لنجاح البرامج هي ٥٥٪ مقارنًا بنسبة معيارية قدرها ٦٠٪ . هل يمكن أن نستنتج من هذا أن عمل الوكالة دون مستوي الأداء المعياري ؟

١- وضع الإفتراضات :

مستوي القياس : مقياس إسمى ثنائي الفئات

النموذج : معاينة عشوائية .

الفرض : ح س = ٦٠و

لقد اخترنا هذا المثال المطابق للمثال السابق بخاصة لنبرز الخلاف في وحدات التحليل . وهنا ندرس وكالة واحدة والعينة عبارة عن عملاء ، إما ناجحين أو راسبين في برامج علاجهم . أما في المثال السابق فكانت الوكالات وليس الأفراد هي التي تم اختيار العينة منها ، وكان المقياس لكل وكالة هو النسبة المئوية للحالات التي نجح فيها البرنامج . لاحظ أننا نحتاج

إلى أي افتراض آخر بالإضافة إلى الفرض الخاص بمجتمع البحث ، حيث من الواضح أن الافتراض الضمني أن توزيع مجتمع البحث ذو منوالين (Bimodal) .

٢- الحصول علي توزيع المعاينة : -

إن توزيع المعاينة سيكون توزيعًا طبيعيًا تقريبًا حيث إن حجم العينة (ن) كبيرٌ .

٣- اختيار مستوي الدلالة والمنطقة الحرجة : -

دعنا نختار ، لمجرد التنويع أن مستوي الدلالة هو عند ٠٢ر ، وأن الاختبار سيكون ذو طرف واحد (One Tailed Test) .

٤ - حساب إحصاءات الإختبار :-

تحسب قيمة الدرجة المعيارية (د) علي النحو التالي :-

$$\text{د} = \frac{(\text{ح}_{\text{س}} - \text{ح}_{\bar{\text{س}}})}{\sqrt{\dfrac{\text{ح}_{\bar{\text{س}}}\,\text{ك}_{\bar{\text{س}}}}{\text{ن}}}} = \frac{\text{٥٥ر} - \text{٦٠ر}}{\sqrt{\dfrac{\text{٦٠ر} \times \text{٤٠ر}}{\text{١٢٥}}}} = \frac{-\text{٠٥ر}}{\text{٠٤٣٨ر}} = -\text{١٤ر١}$$

لاحظ أننا استخدمنا ح $_{\bar{\text{س}}}$ (P_μ) و ك$_{\bar{\text{س}}}$ (q_μ) في المقام بدلاً عن ح$_{\text{س}}$ (P_s) و ك$_{\text{س}}$ (q_s) . وفي حالة ماإذا أردنا أن نستخدم القيمة التائية بدلاً من القيمة المعيارية ، يجب أن نلاحظ أنه طالما توجد قيمة

ح س (P μ) كقيمة إفتراضية ، فإن قيمة الانحراف المعياري لمجتمع البحث (عـ) تحدّد بالمعادلة التالية :-

$$ع = \sqrt{ح_{س} \; ك_{س}}$$

$$\sigma = \sqrt{P\mu \, q\mu}$$

ومن ثم يكون الانحراف المعياري (عـ) معلومًا ، أو علي الأصح يكون مفترضًا .

٥-اتخاذ القرار :-

يمكن أن نري (بعد الإطلاع علي جدول التوزيع الطبيعي) أن القيمة المعيارية (د) التي تبلغ (-١٤و١) أو أقل سوف تحدث تقريبًا في ١٣٪ من الحالات بالصدفة ، في حالة ماإذا كانت الإفتراضات صحيحة ٠ إذن لن نرفض الفرض عند مستوي الدلاله ٠٢, ٠وعلى أساس الدليل الذي بين أيدينا لايمكن أن نحكم بأن مستوي أداء الوكالة أقل من مستوي الأداء المعياري ٠

تمرين الفصل الحادي عشر :-

(١)استخدم جداول الأرقام العشوائية في الجدول (ب) بالملحق رقم (٢) لاختيار ١٠ عينات حجم كل منها أربع مفردات من مجتمع بحث يتكون من ٦٥مفردة كما هو معطي في المسألة رقم (١) بالفصل الرابع ٠أحسب المتوسط الحسابي والانحراف المعياري لكل عينة من العينات العشر ٠ بهذا نكون قد حصلنا علي تقدير تقريبي وقليل التحيز للخطأ المعياري الخاص بالمتوسط الحسابي ٠ كيف تقارن إجابتك بقيمة الخطأ المعياري الذي يُتَحصَّل عليه باستخدام نظرية النهاية المركزية وقيمة الانحراف المعياري

الذي تم حسابه من التمرين رقم (٢) في الفصل السادس ؟

(٢) برهن على أن توزيع المعاينة لمتوسط قذف ثلاث زهرات من النرد هو كما موضح بالشكل رقم (١١ - ٥) ؟

(٣) إذا كانت لدينا عينة ما ، حجمها ٥٠ مفردة ، ومتوسطها الحسابي ١٠,٥ ، وانحرافها المعياري ٢,٢٠ اختبر صحة الفرضية التي مفادها أن المتوسط الحسابي لمجتمع البحث ١٠,٠ باستخدام (أ) الاختبار ذي الطرف الواحد عند مستوي الدلالة ٠٥, .

(ب) الاختبار ذي الطرفين عند مستوي الدلالة ٠١, ؟

أجرِ نفس الاختبارات لعينتين حجم كل منهما ٢٥ مفردة و ١٠٠ مفردة وقارن النتائج التي تحصل عليها ؟

[الإجابة عندما يكون حجم العينة ٥٠، تصبح القيمة التائيه ١,٥٩ ، وهنا نستطيع رفض الفرضية في حالة السؤالين (أ) و (ب)] .

(٤) يفترض أن المتوسط الحسابي للدخل السنوي لعمال مصنع لتجميع قطع منتجات صناعية يبلغ ٩٠٠٠ دولارٍ ، بانحراف معياري قدرة ٩٠٠ دولارٍ. فإذا كنا نشك في أن العمال الذين لديهم اهتمامات في نشاطات الاتحادات العمالية يحصلون علي مداخيل أعلي من متوسط الدخل السنوي . وإذا اخترنا عينة عشوائية من خمسة عشر عاملاً من هؤلاء العمال أصحاب النشاطات في إتحادات العمال ، وكان متوسط دخلهم السنوي ٩٢٠٠ دولارٍ، بانحراف معياري قدره ١٠٠٠ دولارٍ . هل نستطيع أن نستنتج أن هؤلاء العمال النشطين يحصلون علي مداخيل عالية وذات دلالات إحصائية ؟ [استخدم مستوي الدلالة البالغ قدره ٠١,] .

الإجابة د = ٢,٠٥ : وعليه نفشل في رفض الفرض المشار إليـــه في السؤال .

(٥) إذا سحبنا عينة حجمها ٢٠٠شخص من الذين يحق لهم التصويت ، ووجدنا أن من كل اثنين من المرشحين لشغل منصب معين يحصل المرشح (أ) على ٤٥٪ من جملة أصوات العينة . هل نكون محقين إذا قلنا أن المرشح (أ) سيفوز ؟ استخدم مستوي دلالة ٠,٠٥ .اذكر كل الافتراضات التي ينبغي أن نعتمد عليها في حل هذه المسألة ؟

[الإجابة د = ١,١٢] .

(٦) مع أن الشخص لايستطيع حقيقة استخدام التوزيع الاحتمالي الطبيعي المرتبط بالتناسب مالم يكن حجم العينة كبيرًا ، إلا أن المطلوب هنا هو إعادة حل المسألة رقم (٤) بالفصل العاشر كأنها مسألة تختبر فيها الفرض الصفري حول ح $_{س}$(Pµ) ماهو الفرض الصفري (H_O) ؟

قارن الاحتمال المرتبط بالدرجة المعيارية (عـ) مع الاحتمال الصحيح الذي قمت بحسابه من توزيع ذي الحدين ؟

(٧) نفترض أنه قد تم اختيار معيار لاختبار يقيس احتياجات التوافق بين طلبة جميع كليات القطر . فإذا وجدنا أن ٥٠٪ من جميع طلبة الكليات قد حصلوا على ٢٦ درجة فأكثر [ارتفاع الدرجات يعني ارتفاع احتياجات التوافق] . فإذا اعتقد الباحث الإجتماعي بأن احتياجات التوافق تميل إلى الإرتفاع عمومًا بين الكبار الذين لم يحصلوا على تعليم جامعي ، فاختار بذلك عينة عشوائية من بين هؤلاء الكبار الذين تبلغ أعمارهم ٢٥سنة فأكثر ويقيمون في مجتمع ما. فإذا وجد الباحث النتائج التالية :(أ) أن ٦٧٪ من مجموع ٢٥٧من الكبار الذين لم يحصلوا على تعليم جامعي تبلغ درجاتهم

٢٦ درجه فأكثر (ب) أن ٢٩ ٪ من مجموع ٨٠ شخصاً من الكبار الذين تحصلوا علي تعليم جامعي قد نالوا درجات في المدي ذاته . (أ) هل يمكن للباحث أن يستنتج أن الدرجات التي حصلت عليها أي من المجموعتين أعلي بدلالة إحصائية من الدرجات الخاصة بجميع طلاب الكليات الذين خضعوا للاختبار المعياري ؟ استخدم مستوي دلالة ٠,٠٠١ .

(ب) نفترض أن الباحث الاجتماعي يعلم بالضبط التوزيع الكلي لدرجات طلاب الكليات في الاختبار علي ضوء ماسبق شرحه في هذا الفصل ، أذكر بعض الإجراءات البديلة لاختبار مستوي الدلالة الذي يوضح مدي ابتعاد مجموعتي الكبار عن الدرجات المعيارية لجميع طلاب الكليات ؟ هل تحتاج هذه الإجراءات البديلة لأي إفتراضات إضافية ؟ وضح ذلك ؟.