دراسة تحليلية مقارنة للصيغ المستخدمة في حساب حجم العينة العشوائية

الدكتور شكيب بشماني*

(تاريخ الإيداع 23/ 7 / 2014. قُبِل للنشر في 20 /10 / 2014)

□ ملخّص □

يهدف البحث إلى وضع بعض الصيغ المعبرة عن حجم العينة و توصيفها والمقارنة فيما بينها لتحديد الصيغة الأفضل من بين تلك الصيغ لحساب حجم العينة وإيجاد صيغة معدلة تعبر تعبيراً جيداً عن حجم العينة ، بالإضافة إلى تحديد حدي الاشباع الأول والثاني للصيغ ذات الصلة ووضع معادلات رياضية يمكن من خلالها التنبؤ بحجم العينة مهما بلغ حجم المجتمع .

توصل الباحث من خلال دراسته إلى النتائج الآتية:

- إن النتائج كانت متطابقة بالنسبة للصيغ ذات الصلة بحجم المجتمع وحجم العينة عند توحيد شروطها.
- لم يزدد حجم العينة معنوياً مع زيادة حجم المجتمع عند حد الاشباع الأول .
- لا توجد فروق معنوية بين حجمي العينة وفقا لحجم المجتمع عند حدي الاشباع .
- توجد فروق معنوية بين حجم العينة ومتوسط إجمالي الفحص وفقاً لحجم المجتمع عند حدي الاشباع .
- حصلنا على نماذج رياضية للعلاقة بين حجم المجتمع وحجم العينة وكذلك بين حجم المجتمع ومتوسط اجمالي الفحص .
- توصلنا إلى وضع جدول شامل يعطينا حجم العينة المقابل لحجم المجتمع يمكن أن يكون في متناول الباحثين للاستفادة منه والاستغناء عن تطبيق الصيغ طالما هو يعتمد عليها بالأصل عند شروط معينة .

الكلمات مفتاحية: حجم المجتمع، حجم العينة العشوائية، صيغ حساب حجم العينة , حد الاشباع .

*مدرّس–قسم الإحصاء والبرمجة–كلية الاقتصاد–جامعة تشرين–اللاذقية–سورية.

Tishreen University Journal for Research and Scientific Studies -Economic and Legal Sciences Series Vol. (36) No. (5) 2014

Comparative Analysis of Formulas Used to Calculate the Size of the Random Sample

Dr. Shakeeb Bishmani*

(Received 23 / 7 / 2014. Accepted 20/ 10/ 2014)

☐ ABSTRACT ☐

The research aims to develop some formulas of sample size and characterization and comparison among themselves to determine the best formula of formulas to calculate the sample size and the creation of a modified reflected well on the sample size, in addition to specifying individual gratification I and II for the relevant formulas and mathematical equations can predict the sample size, however the size of the community. The researcher through the study the following results:

1. The results were identical to the formula related to the size and the sample size when consolidation requirements.
2. Sample size did not increase with increasing size of the moral community at first gratification.
3. No moral differences between sample volume according to the size of the community when individual gratification.
4. Moral differences exist between sample size and average total inspection according to the size of the community when individual gratification.
5. We got a mathematical models of the relationship between size and the sample size and the size of the community and the average total inspection.
6. We have developed a comprehensive table gives sample size corresponding to the size of the community can be accessible to researchers to take advantage of it and apply the formulas as long as it originally relied upon certain conditions.

Keywords: Population Size, Random Sample Size; Formulas of Computing of Random Sample Size, Gratification.

مقدمة:

*Assistant Professor, Department of Statistics & Programming; Faculty of Economics, Tishreen University, Lattakia, Syria.

كل باحث يدرك ما لحجم العينة العشوائية من أهمية بالغة في البحوث الإحصائية وغير الإحصائية ؛ لذلك لا بد من إيلاء هذا الحجم أهمية بالغة عند تطبيق مبدأ العشوائية في السحب . بناء عليه هناك صيغ عديدة في حساب حجم العينة ؛ لذلك يجب الانتباه إلى الصيغة المطبقة في حسابه عند شروط معينة من حيث حجم المجتمع والخطأ المسموح به والقيمة الاحتمالية وغيرها من الشروط ذات الصلة .

إن الوصول إلى حجم عينة مناسب والسحب بطريقة غير متحيزة تكون فيها العينة معبرة تعبيراً صادقاً للمجتمع هاجس كل إحصائي أو باحث يسعى إلى الدقة في الوصول إلى نتائج أبحاثه بحيث يستطيع بعد ذلك تعميم نتائج أبحاثه على المجتمع والتي حصل عليها من خلال العينة ، والاستفادة منها في اتخاذ قرارات سليمة رشيدة ، كل هذا يعتمد على استخدام صيغة مناسبة لحساب حجم العينة سواء كان المجتمع صغيراً أم كبيراً وسواء كان التطبيق في الأبحاث العامة أم أبحاث الجودة العائدة لمخرجات خطوط الإنتاج أو بعد الإنتاج أو حتى بعد البيع. يسعى الباحث في هذا البحث إلى توصيف الصيغ ذات الصلة بحجم العينة وتحديد علاقة حجم العينة بحجم المجتمع وإيجاد حدي الاشباع لحجم العينة ، بالإضافة إلى وضع جدول شامل لحجم العينة المقابل لحجم المجتمع يكون في متناول جميع الباحثين مضافاً إليه عمود يخص متوسط إجمالي الفحص المستخدم في مجالات وأبحاث مراقبة الجودة ، ثم وضع معادلات رياضية تمكننا من معرفة سلوك هذه الصيغ والتنبؤ بقيم حجم العينة عند أي حجم للمجتمع .

مشكلة البحث:

يكثر استخدام صيغ حساب حجم العينة العشوائية في الأبحاث المختلفة وبخاصة أبحاث الدراسات العليا ، إذ هناك عدد كبير من هذه الصيغ منها يتضمن حجم المجتمع في الصيغة ومنها لا يتضمن ذلك . والأهم من ذلك أنه لا يوجد توصيف وتحليل واضح لتحديد الصيغ المناسبة في حساب حجم العينة وعلاقتها بحجم المجتمع، بالإضافة إلى غياب الآلية الواضحة لسحب مفردات العينة باستخدام الصيغ الإحصائية والرياضية ، عدا ذلك عدم وجود نموذج رياضي يوضح الاتجاه العام وسلوك هذه الصيغ عند شروط مشتركة لكل الصيغ ، وعدم تحديد درجة الإشباع لهذه الصيغ وعدم وجود جدول جاهز يعطي حجم العينة المقابل لحجم المجتمع في لحظته يكون في متناول الباحثين. هناك أخطاء عديدة يقع فيها الباحثون وبخاصة طلاب الدراسات العليا عند تحديد حجم العينة في أبحاثهم لذلك يجب وضع آلية واضحة تناسب الباحثين في تحديد حجم العينة المستخدمة في بحوثهم سواء التي تعتمد على بيانات مستمرة أو على بيانات منقطعة .

أهمية البحث وأهدافه:

الأهمية:

تتجلى أهمية البحث في النقاط التالية:

1- أهمية دراسة العلاقة بين حجم المجتمع الإحصائي وحجم العينة العشوائية كونها تضع الخطوط الموضوعية لأي بحث إحصائي .

2- أهمية أن تكون العينة ممثلة تمثيلاً صادقاً للمجتمع من خلال سحب عدد مفردات كافيه تعكس الخصائص الرئيسية والحقيقية لهذا المجتمع.

3- وضع الخطوط العريضة في آلية سحب مفردات العينة باستخدام الصيغ الإحصائية والرياضية المناسبة .

4- وضع نموذج رياضي لحساب حجم العينة أياً كان حجم المجتمع وذلك بتعديل وتطوير وتعميم الصيغ المستخدمة في حساب حجم العينة وبخاصة عندما يصل حجم المجتمع إلى درجة الإشباع ، إذ يثبت عندها حجم العينة مهما كبر حجم المجتمع .

الأهداف:

ويهدف البحث إلى:

1- وضع بعض الصيغ المعبرة عن حجم العينة و توصيفها والمقارنة فيما بينها .

2- تحديد الصيغة الأفضل من بين تلك الصيغ لحساب حجم العينة .

3- إيجاد صيغة معدلة تعبر تعبيراً جيداً عن حجم العينة وتمثل مفرداتها تمثيلاً جيداً لحجم المجتمع من حيث النسبة التي يمثلها حجم العينة من حجم المجتمع وإمكانية التنبؤ بحجم العينة مهما تغير حجم المجتمع ، كبُر أم صغُر.

متغيرات البحث:

1- المتغير المستقل: حجم المجتمع .

2- المتغير التابع: حجم العينة.

فرضيات البحث:

1- لا توجد فروق ذات دلالة إحصائية بين الصيغ المختلفة لحساب حجم العينة عند توحيد الشروط للصيغ .

2- لا يزداد حجم العينة مع زيادة حجم المجتمع عند حد إشباع معين .

3-لا توجد فروق ذات دلالة إحصائية بين حجمي العينة عند حدي الاشباع وفقاً لحجم المجتمع .

4-لا توجد فروق ذات دلالة إحصائية بين حجم العينة ومتوسط إجمالي الفحص وفقاً لحجم المجتمع عند حدي الإشباع الأول والثاني .

منهجية البحث:

اعتمد البحث على المنهج الوصفي التحليلي في توصيف الصيغ العائدة لحساب حجم العينة من حجم المجتمع، كما اعتمد على المنهج الاستنباطي بالاستناد إلى العلاقات الرياضية والإحصائية التي تم التوصل إليها وإمكانية التنبؤ بحجم العينة .

مجتمع وعينة البحث:

هناك عدد لا بأس به من الصيغ الإحصائية والرياضية المستخدمة من الباحثين المختلفين في تحديد حجم العينة ولا يمكن حصرها بالتحديد . سنأخذ بعض هذه الصيغ وعددها سبع صيغ وهي الأكثر تطبيقاً في الأبحاث وبخاصة أبحاث الدراسات العليا ، أما حجم المجتمع المطبق عليه صيغ حجم العينة فهو يبدأ من الحجم 30 مفردة إلى 30000 كمرحلة أولى في تحديد حجم العينة ، أما المرحلة الثانية فهي تتجاوز حجم 30000 مفردة لتصل إلى اللانهاية . سنطبق هذه الصيغ السبع لنحصل منها على حجم العينة المناسب لحجم المجتمع والمقارنة فيما بينها ، ثم الوصول إلى صيغة متوسط إجمالي الفحص (ATI) المستخدم في مراقبة الجودة .

حدود البحث :

-الحدود المكانية : لا حدود مكانية للبحث إلا حدود المجتمع الافتراضي ؛ لأنه بحث إحصائي رياضي تحليلي.

-الحدود الزمانية : تم إعداد البحث في الفترة الواقعة من 6/5 /2013 ولغاية 21 /7/ 2014 م .

الدراسات السابقة:

دراسة (Bartlett , 2001) بيّن العلاقة بين حجم المجتمع وحجم العينة لنوعين من البيانات ، النوع الأول : البيانات المنقطعة وفقاً لهامش خطأ (0.05) وقيمة P الاحتمالية (0.50) ، أما مستويات الدلالة فقد درسها عند ثلاثة مستويات تتمثل بـ (0.10 , 0.05 , 0.01) ووضع الدرجات المعيارية المقابلة لها ، والنوع الثاني : كان يخص البيانات المستمرة وبهامش خطأ (0.03) وعند مستويات دلالة ثلاثة كما في النوع الأول ووضع كذلك الدرجات المعيارية المقابلة لها . توصّل من خلال ما تقدم إلى وضع جدول بنوعي البيانات الوارد ذكرها آنفاً ، حصل من خلاله على حجم العينة المناسب لكل حجم مجتمع ولكل نوع من البيانات بدءاً من الحجم (100) للمجتمع وانتهاءً بالحجم (10000) وقد لخص هذا الجدول بـ (15) قيمة لحجم المجتمع يقابله (15) قيمة لحجم العينة لنوعي البيانات .[6] استخدم صيغة كوكران Cochran في إيجاد قيم حجم العينة لنوعي البيانات سواء المستمرة أم المنقطعة والتي تأخذ الصيغة التالية للبيانات المستمرة :

$$n_0 = (\frac{t^2 * s^2}{d^2} \qquad \text{(A)}$$

أما صيغة كوكران للبيانات المنقطعة و الفئوية فتأخذ الصيغة التالية :

$$\bar{n}_0 = \frac{t^2 * (p)(q)}{d^2} \qquad \text{(B)}$$

tهنا تقابل z=1.96 , p=q= 0.50 ,

ثم وضع صيغة لحجم العينة المصحح لكلا النوعين من البيانات ليعتبره حجم العينة النهائي كما في الصيغة التالية :

$$\underline{n_1} = \frac{n_0}{(1 + \frac{n_0}{N})} \qquad \text{(C)}$$

نشير إلى أن الدراسة الحالية تميزت عن الدراسة السابقة بأنه تم إعداد جدول لنوعي البيانات من الحجم (30) إلى اللانهاية مبينة حدي الإشباع الأول والثاني ، وكذلك طريقة التحليل إذ لم تعتمد على التحليل الإحصائي فقط بل اعتمدت التحليل الرياضي أيضاً ووضع نماذج رياضية تبين سلوك هذه الصيغ عند توحيد الشروط ، بالإضافة إلى تعديل النماذج بدمج متوسط إجمالي الفحص إلى النموذج الرياضي المعبر عن حجم العينة ليعبر عن حجم العينة النهائي ، وبخاصة عند استخدامه في بحوث مراقبة الجودة .

المجتمع والعينة:

أولاً: مفهوم المجتمع الإحصائي: هو أي مجموعة من العناصر تشترك بخاصة معينة أو أكثر وتكون خاضعة للدراسة الإحصائية .[1] يرمز لحجم المجتمع الإحصائي – الذي يعني عدد مفرداته – بالرمز N .

ثانياً: مفهوم العينة العشوائية : هي مجموعة جزئية من وحدات المجتمع بحيث يكون حجم العينة ممثلاً لها والمعلومات المتوفرة في العينة هي تلك المتوفرة في المجتمع [3] . تسحب العينة عشوائياً وتدرس خصائص

وحداتها لتعمم مميزات تلك الخصائص على المجتمع الإحصائي الكلي.[1] يرمز لحجم العينة – الذي يعني عدد مفردات العينة المسحوبة من المجتمع – بالرمز n وهو مرتبط بمقدار الدقة والثقة التي يرغب الباحث بتحقيقها .

ثالثاً: حدي الإشباع: [الباحث]

–حد الإشباع الأول : وهو الحد الذي لا يزيد عنده حجم العينة مع زيادة حجم المجتمع زيادة معنوية ، وقد تكون الزيادة طفيفة أو شبه معدومة بالمقارنة مع زيادة حجم المجتمع .

–حد الإشباع الثاني : وهو الحد الذي لا يزيد عنده حجم العينة أبداً مع زيادة حجم المجتمع .

رابعاً: صيغ حساب حجم العينة العشوائية :

هناك صيغ عديدة لحساب حجم العينة ولا يمكن حصرها ، لكن يمكن التركيز على بعض منها وهي الصيغ الأكثر استخداماً في بحوث الدراسات العليا المتمثلة بالصيغ السبع التالية :

$$n=\frac{p(1-p)}{(E\div Z)+[p(1-p)\div N]} \tag{1}$$

وتسمى بمعادلة هيربرت أركن

حيث :

Nحجم العينة و N حجم المجتمع .

P قيمة احتمالية تتراوح قيمتها بين الصفر والواحد وتأخذ قيمة 0.50 أينما وجدت لتثبيت الشروط ولأننا لا نعرف تقدير P ومعظم الأحيان يكون غير متاح لذلك نأخذ القيمة العظمى وهي $P(1-P) = 0.25$.

E نسبة الخطأ المسموح به وثبتناها هنا عند 0.05 .

Z الدرجة المعيارية وتساوي 1.96 عند معامل ثقة 0.95 .

$$n=\frac{N.Z^2.S^2}{N.d^2+Z^2.S^2} \tag{2}$$

وهي في حالة السحب بدون إعادة .

d: الدقة المطلوبة .

S^2.: تباين العينة كتقدير لتباين المجتمع وهنا يجب أن يكون تباين العينة معلوماً .

$$n=\frac{\left(\frac{z}{d}\right)^2\times(P)^2}{1+\frac{1}{N}\left[\left(\frac{z}{d}\right)^2\times(P)^2-1\right]} \tag{3}$$

تسمى معادلة ريتشارد جيجر .

$$n=\frac{p.q.z^2}{E^2} \tag{4}$$

$$n = \frac{z_{1-\frac{\alpha}{2}}^{2} . s^{2}}{E^{2}} \qquad (5)$$

تسمى المعادلتان (4) و (5) بمعادلتي الحد الأدنى لحجم العينة في حالة السحب مع الإعادة ولا يدخل في الصيغتين رمز حجم المجتمع لذلك سيكون لهما معاملة بعيدة عن العلاقة الجدلية بين حجم المجتمع وحجم العينة ، إذ تعتمدان بالدرجة الأولى على خطأ التقدير والتباين المقدر ، وهنا يجب أن يكون تباين العينة معلوماً كتقدير لتباين المجتمع . فالصيغة (4) تستخدم عندما يكون السحب مع الإعادة والمجتمع كبير جداً [5] ، وكذلك تستخدم عندما تكون البيانات منقطعة أو حتى فئوية ، أما الصيغة رقم (5) فهي تستخدم عندما تكون البيانات مستمرة وهما صيغتا كوكران لنوعي البيانات المنقطعة والمستمرة .

$$n = \frac{N \times p(1-p)}{\left[\left[N-1 \times \left(d^{2} \div z^{2}\right)\right] + p(1-p)\right]} \qquad (6)$$

تسمى معادلة ستيفن ثامبسون .

حيث P(1−P) تباين نسبة المجتمع فإذا كانت P غير معلومة فهي تعامل أيضاً كقيمة عظمى لـ P(1−P) أي 0.25 = (0.50) . (0.50) وهي تعامل نفس المعاملة في أية صيغة ترد فيها إذا كانت P غير معلومة.

q = (1−P) النسبة المتبقية للخاصية p .

$$n = \frac{N}{\left[\left(S^{2} \times (N-1)\right) \div pq\right] + 1} \qquad (7)$$

تسمى بمعادلة روبرت ماسون .

النتائج والمناقشة:

بالنظر إلى الصيغ السابقة نجد أن هناك صيغاً تخص نسبة المجتمع وتتغير مباشرة بتغير حجم المجتمع وهي الصيغ ذوات الأرقام (1,3,6,7) وهي من سيتم التركيز عليها في البحث ، وكذلك هناك صيغ تخص التباين المقدر للمجتمع وتتغير مباشرة بتغير حجم المجتمع وهي الصيغة رقم (2) ، وصيغة تخص النسبة وللبيانات المنقطعة وليس لها علاقة مباشرة بحجم المجتمع وهي الصيغة رقم (4) وصيغة تخص التباين المقدر وللبيانات المستمرة وليس لها علاقة مباشرة بحجم المجتمع وهي الصيغة رقم (5) . نلاحظ من علاقتي كوكران (4) و (5) أنهما لا يتبعان مباشرة إلى حجم المجتمع ولكن لاستخدامهما لا بد من تعديل حجم العينة الذي نحصل عليه منهما لأنه يعطي حجم عينة

أولي باستخدام الصيغة (C) لـ Bartlett التي تعطينا حجم العينة المصحح أو النهائي والتي تحتوي في صيغتها المصححة على حجم المجتمع.

عند تثبيت جميع الشروط لجميع الصيغ أو المعادلات السابقة التي لها صلة مباشرة بحجم المجتمع نحصل على نفس النتائج لحجم العينة كما هو مبين في الجدول رقم (1) ما عدا الصيغتين (4) و (5) اللتين لا يمكن تضمينهما في صيغ الجدول لأنهما لا يحتويان على حجم المجتمع ؛ إلا أن الصيغة رقم (4) تعطي النتيجة نفسها التي نحصل عليها من الصيغ المتبقية عند حد الإشباع الثاني وهو (n=385) عندما توحد شروطها مع شروط الصيغ السابقة ، بينما الصيغة (5) تعطينا القيمة (385) لحجم العينة ،أي ما يقابل حد الإشباع الثاني لبقية الصيغ عندما يكون (S=0.25 , E = 0.025 , Z=1.96) ، أما إذا كان تباين المجتمع مجهولاً فتتحول الصيغة رقم (5) إلى الصيغة رقم (5-A) التالية[4] :

$$\boxed{n = (\frac{Z_{\frac{\alpha}{2}}}{2E})^2} \qquad (5\text{-}A)$$

وتعطي عندئذٍ نفس نتيجة حجم العينة للصيغ السابقة عند حد الإشباع الثاني بشرط توافر الشروط نفسها ويكون هذا الحجم مساوياً أيضاً (385) مفردة .

طالما العلاقة (C) لـ Bartlett يدخل فيها حجم المجتمع سنبحث عن حجم العينة المقابل لحدي الاشباع مقارنة مع الصيغ ذات الصلة التي أعطتنا الجدول (1) بتطبيق العلاقة (C) فنجد :

–حد الاشباع الأول:

$$\underline{n_1} = \frac{380}{(1 + \frac{380}{30000})} = 376$$

وهو يختلف عن حد الإشباع الأول للصيغ السابقة قليلاً أي أخفض بأربع درجات ، والسبب يعود إلى أن هذه الصيغة هي الصيغة المصححة لحجم العينة التي تضمنت حجم العينة الأولي n_0 واستخدامه للحصول على حجم العينة النهائي (المصحح) ،أي $\underline{n_1}$ لكن لو طبقنا الصيغ الخمس السابقة سنحصل على النتيجة 380 وهي القيمة نفسها لحد الإشباع الأول . لم نستطع تطبيق صيغتي كوكران غير المصححتين (4) و(5) لأنهما لا تحتويان على حجم المجتمع لذلك لجأ Bartlett إلى تعديل صيغتي كوكران من وجهة نظره ليحصل على الصيغة C . علماً أن حد الإشباع الأول تم تحديده من خلال بيانات الجدول التي حصلنا عليها بتطبيق الصيغ ذات الصلة والتي أعطت نتائج لحجم العينة لا تختلف كثيرا بدءاً من (30000) مفردة ونلاحظ ذلك من خلال الجدول (1) الآتي أن حجم العينة قد ازداد من (380) مفردة المقابل لحجم المجتمع (30000) مفردة إلى (381) مفردة المقابل لحجم المجتمع (40000) مفردة بمعدل زيادة مقداره (0.0026) مفردة لحجم العينة ، أي 0.26 % وهذا ضئيل جداً بالمقارنة مع حجم المجتمع الذي ازداد (10000) مفردة ، أي بمعدل زيادة (0.3333) يعني 33.33 %.

–حد الاشباع الثاني :

$$\underline{n_1} = \frac{385}{(1+\frac{385}{1000000})} = 385$$

هذا يعني أن حد الاشباع الثاني لحجم المجتمع هو نفسه (1000000) مفردة عند علاقة بارتليت التي يقصد بها حجم العينة المصحح أو النهائي المقابل هو(385) مفردة وبذلك تشترك جميع الصيغ السابقة الذكر عند حد الاشباع الثاني ، أي لا يوجد اختلاف بينها عند هذا الحد . نلاحظ من الجدول (1) أيضاً أن حجم العينة يثبت عند (384) مفردة بدءاً من حجم المجتمع (150000) مفردة ليصبح (385) مفردة عند حجم المجتمع (1000000) مفردة ويثبت عند هذا الحجم مهما زاد حجم المجتمع .

الجدول رقم (1) العلاقة بين حجم المجتمع وحجم العينة ومتوسط إجمالي الفحص ((جزء بسيط من الجدول الإجمالي الموجود في الملحق (1)) مع تثبيت الشروط التالية : (α=0.05 , E = 0.05 , Z = 1.96)

ATI=n+(1-p)(N-n)	P=n/N	n	N
29	0.93333333	28	30
38	0.925	37	40
46	0.9	45	50
53	0.86666667	52	60
61	0.85714286	60	70
68	0.8375	67	80
74	0.81111111	73	90
81	0.8	80	100
315	0.278	278	1000
590	0.0714	357	5000
827	0.03894737	370	9500
852	0.037	370	10000
904	0.03381818	372	11000
955	0.03108333	373	12000
1006	0.02876923	374	13000
1056	0.02671429	374	14000
1107	0.025	375	15000
1158	0.0235	376	16000
1208	0.02211765	376	17000
1259	0.02094444	377	18000
1309	0.01984211	377	19000
1359	0.01885	377	20000
1611	0.01516	379	25000
1861	0.01266667	380	حد الإشباع الأول 30000
2362	0.009525	381	40000
7865	0.00256	384	150000
45365	0.0004267	384	900000
50366	0.000385	385	1000000 حد الإشباع الثاني
∞	0	385	∞

المصدر : من إعداد الباحث بالاعتماد على صيغ حساب حجم العينة وقيم حجم العينة مقربة إلى أعلى رقم صحيح.

يمكن إعداد جدول آخر بتغيير الشروط حسبما يراها الباحث كأن يكون (α=0.01) فنحصل على جدول جديد يختلف عن الجدول السابق .

يمكن برمجة هذه الصيغ على MS EXCEL للحصول على أية قيمة لكن لن تكون مقربة إلى أعلى رقم صحيح وهنا يتوجب على الباحث أن يأخذ بعين الاعتبار هذا التقريب في أي بحث يطبق فيه إحدى هذه الصيغ .

–اختبار الفرضية الأولى : لا توجد فروق ذات دلالة إحصائية بين الصيغ المختلفة لحساب حجم العينة عند توحيد الشروط .

نلاحظ من الجدول (1) السابق أن جميع المعادلات والصيغ أعطت النتائج نفسها لحجم العينة ، أي النتائج كانت متطابقة وهذا يستدعينا القول بأنه لا توجد فروق بين الصيغ المختلفة وبالتالي قبول الفرضية الأولى بكل بساطة .

–اختبار الفرضية الثانية: لايزداد حجم العينة مع زيادة حجم المجتمع عند حد إشباع معين (حد الاشباع الأول)

وجدنا أن حد الإشباع الأول لهذه الصيغ هو عندما يكون حجم المجتمع (30000) مفردة يقابله حجم للعينة مساوياً (380) مفردة عندئذٍ ستكون الزيادة طفيفة جداً لحجم العينة مهما زاد حجم المجتمع لدرجة أن الصيغ جميعها تعطي حجماً للعينة مساوياً لـ (385) عندما يصبح حجم المجتمع (1000000) وحتى عندما يصبح حجم المجتمع لانهائي فهي جميعها تعطي حجم العينة نفسه ألا وهو (385) مفردة باستثناء الصيغتين رقم (2) ورقم (6) فهما تعطيان حالة عدم تعيين من الشكل ($\frac{\infty}{\infty}$) فإذا أزلنا حالة عدم التعيين السابقة بتطبيق قاعدة أوبيتال (مشتق البسط على مشتق المقام) [6] فسنحصل على حجم العينة نفسه وهو (385) وهذا هو القصور بعينه لهذه الصيغ فهي تسلك سلوكاً ثابتاً عند حد الاشباع الثاني (1000000) مفردة وسلوكاً شبه ثابت عند حد الإشباع الأول (30000) مفردة .

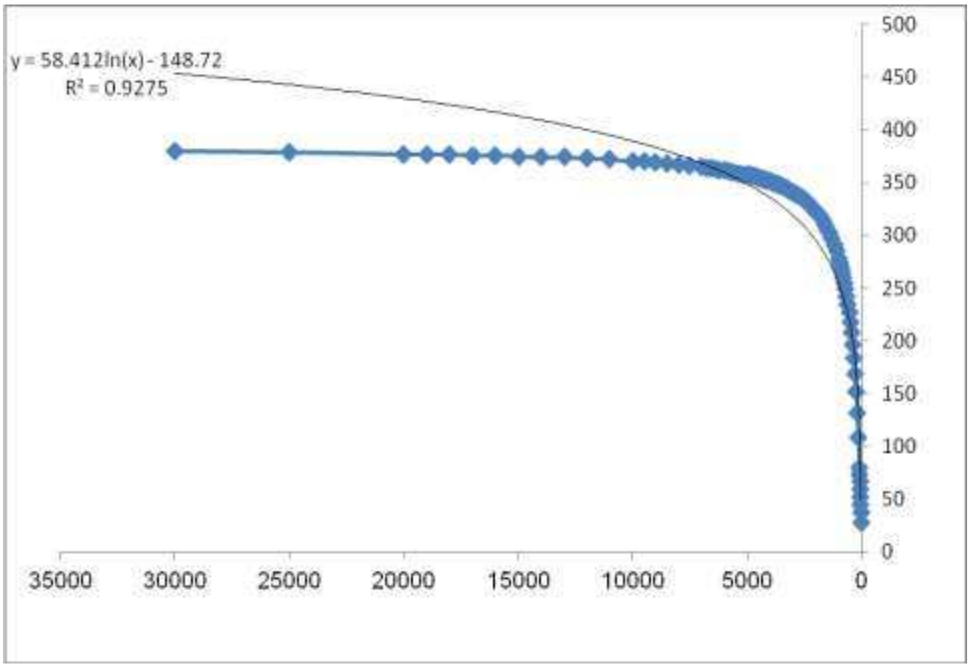
–اختبار الفرضية الثالثة : لا توجد فروق ذات دلالة إحصائية بين حجمي العينة عند حدي الاشباع وفقاً لحجم المجتمع .

أخذنا 94 مفردة كعينة أولى من المرحلة الأولى (أي المجتمع الأول) للجدول (1) إي إلى حد الاشباع الأول فحصلنا على القيم التالية : الوسط الحسابي لحجم العينة يساوي 299.17 والتباين يساوي 9123.03 وكذلك أخذنا 94 مفردة من المرحلة الثانية (المجتمع الثاني) كعينة ثانية فحصلنا على النتائج التالية : الوسط الحسابي لحجم العينة يساوي 383.75 والتباين يساوي 0.4718 . بتطبيق قانون مؤشر الاختبار للفرق بين متوسطي مجتمعين عند مستوى دلالة (α=0.05) حصلنا على قيمة Z الفعلية تساوي (Z=-8.59) بينما قيمة Z الجدولية (Z = 1.96) . بالمقارنة نجد القيمة الفعلية بالقيمة المطلقة أكبر من القيمة الجدولية ؛ لذلك نقول أن هناك فرق معنوي بين المتوسطين الناتجين من تطبيق الصيغ في المجتمعين (المجتمع الأول حتى حد الاشباع الأول والمجتمع الثاني بين جدي الاشباع) .

سنأخذ القيمتين المقابلتين لحدي الاشباع وسنعتبرهما ممثلين لمتوسطي العينتين نظراً لأنهما ناتجتين عن تطبيق صيغ حجم العينة وبالاعتماد على تبايني العينتين السابقتين وعدد مفردات لكل عينة (94) مفردة ، وسنختبر الفرق بين المتوسطين عند مستوى الدلالة السابق نفسه (0.05) لنحصل على القيمة الفعلية التالية (Z= - 0.5075) مقابل القيمة الجدولية (Z = 1.96) وبالمقارنة سنجد القيمة الجدولية أكبر من القيمة الفعلية ؛ لذلك نقول لا يوجد فرق جوهري بين القيمتين ، أي بين حدي الاشباع الأول والثاني وبالتالي ستكون الزيادة هنا في حجم المجتمع بين حدي الاشباع تقابلها زيادة غير معنوية في حجم العينة .

–<u>اختبار الفرضية الرابعة :</u> لا توجد فروق ذات دلالة إحصائية بين حجم العينة ومتوسط إجمالي الفحص وفقاً لحجم العينة عند حدي الاشباع الأول والثاني .

سنجد الاتجاه العام لهذه الصيغ إلى حد الإشباع الأول (30000) مفردة وهو كافٍ للحكم على اتجاه هذه الصيغ مع ازدياد حجم المجتمع إلى هذا الحد . بالاعتماد على برنامج Excel حصلنا على الشكل البياني رقم (1) التالي:



الشكل رقم (1) يبين الاتجاه العام للعلاقة بين حجم العينة وحجم المجتمع

المصدر : من إعداد الباحث بالاعتماد على بيانات حجم المجتمع والعينة باستخدام الصيغ المختلفة وبرنامج Excel

درسنا العلاقة بين المتغيرين (حجم المجتمع وحجم العينة) فوجدنا ومن خلال الشكل البياني (1) السابق أن أفضل المعادلات التي تمثل هذه العلاقة هي المعادلة اللوغاريتمية التالية :

$$n = 58.41\ln(N) - 148.7 \qquad (8)$$

إن معامل التحديد لهذه المعادلة يساوي : $R^2 = 0.927$

أي أن حجم المجتمع في هذه المعادلة يفسر ما مقداره (92.7 %) من التباين الحاصل في حجم العينة .

من خلال ما تقدم نلاحظ أن بالإمكان وضع نموذج رياضي للعلاقة بين حجم المجتمع وحجم العينة حتى حد الإشباع الأول .

وبالاعتماد على برنامج spss18 نحصل على الجدول رقم (2) الذي يؤيد صحة ما حصلنا عليه في برنامج Excel لقيمة معامل التحديد :

الجدول رقم (2) يظهر معامل التحديد وخطأ التقدير .

R	R Square	Adjusted R Square	Std. Error of the Estimate
------	.928	.927	25.855

The independent variable is population size.

وبإعداد جدول تحليل ANOVA نحصل على الجدول رقم (3) التالي :

الجدول رقم (3) جدولANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	786940.770	1	786940.770	1177.203	.000
Residual	61500.507	92	668.484		
Total	848441.277	93			

The independent variable is population size.

من خلال جدول ANOVA نجد أن قيمة sig = 0 أصغر من مستوى الدلالة α=0.05 وبالتالي نقول بأن هناك أثراً ذا دلالة إحصائية لحجم المجتمع في حجم العينة . نلاحظ أيضاً ثبات صلاحية النموذج استناداً إلى ارتفاع قيمة F) المحسوبة والبالغة (1177.203) عن قيمتها الجدولية عند مستوى دلالة (α=0.05) ودرجات حرية (92,1) والبالغة (2.53) .

ولإيجاد ثوابت المعادلة واختبارها وفق برنامج spss نحصل على الجدول رقم (4) التالي :

جدول رقم (4) ثوابت معادلة التمثيل واختبارها

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
Ln (population size)	58.412	1.702	.963	34.310	.000
(Constant)	-148.716-	13.324		-11.162-	.000

نلاحظ من الجدول (4) أن المعادلة الممثلة للعلاقة بين حجم المجتمع وحجم العينة وفق الصيغ ذات الصلة يمكن تمثيلها بمعادلة لوغاريتمية لتأخذ الشكل التالي :

$$n=58.412\ln(N) - 148.716 \qquad (8\text{-}A)$$

بالنظر إلى الجدول نجد أن قيمة sig المقابلة للثابت تساوي صفر بينما قيمة مستوى الدلالة α=0.05 بالمقارنة نجد أن قيمة sig أصغر من قيمة α هذا يعني أن قيمة الثابت معنوية .

وهي تأخذ نفس قيمة المعادلة (8) مع الفرق بالتقريب فقط .

يفضل استخدام المعادلة (8) ولنكون أكثر دقة يفضل استخدام العلاقة (8-A) عندما يكون حجم المجتمع 1000 مفردة وما فوق نظراً للنسبة العالية التي يأخذها حجم العينة من حجم المجتمع إذا كان حجم المجتمع أقل من 1000 مفردة .

نلاحظ أيضاً من الجدول رقم (1) نسبة حجم العينة من حجم المجتمع اعتمادا على الصيغ السابقة ، لكن لم نكتفِ بوضع نموذج رياضي للعلاقة بين المتغيرين السابقين بل اعتمدنا على معادلة متوسط إجمالي الفحص أو السحب المستخدم في مراقبة الجودة ؛ إذ يفيدنا هذا المتوسط في حالة كون المجتمع كبيراً والمفردات تحتاج لأكبر فحص ممكن كونها تتمتع بخاصية الحساسية وضرورة الدقة في السحب والفحص.

تعطى علاقة متوسط إجمالي الفحص بالعلاقة التالية:[2]

$$ATI=n + (1-Pa)(N-n) \qquad (9)$$

حيث:

nحجم العينة.

N : حجم المجتمع .

Pa : احتمال الثقة .

بتطبيق العلاقة (9) حصلنا على النتائج المبينة في الجدول رقم (1) العمود رقم (4) .

نلاحظ عند حد الاشباع الأول أن حجم العينة بالصيغ السابقة كان يساوي (380) مفردة وأصبح بتطبيق صيغة متوسط إجمالي الفحص (1861) مفردة وهناك فرق واضح وملموس لا داعي لاختباره . أما عند حد الاشباع الثاني فقد كان حجم العينة مساوياً (385) مفردة بتطبيق جميع الصيغ الصيغ السابقة بعد إزالة حالات عدم التعيين ، أما بتطبيق صيغة متوسط إجمالي الفحص فسنجد أن حجم العينة يساوي (50366) مفردة وهناك فرق كبير في حجم العينة بين الطريقتين ؛ لذلك يجب الانتباه عند اختيارنا للصيغة ، فالصيغ السابقة ذات الصلة بحجم العينة وحجم المجتمع لا تلبي حجم العينة المطلوب عند كبر حجم المجتمع وبخاصة بعد حد الاشباع الأول ، أما صيغة متوسط إجمالي الفحص فهي تغطي أكثر بكثير من الصيغ السابقة .إذاً هناك فروق معنوية بين قيم حجم العينة الناتجة عن تطبيق صيغ حجم العينة و قيم حجم متوسط اجمالي الفحص عند حدي الاشباع ، وهذا فرق واضح لا يحتاج إلى اختبار .

إن متوسط إجمالي الفحص هو طريقة لتقويم خطة العينة وهو الكمية المسحوبة أو المفحوصة [2] .

بناء على ما تقدم تصبح المعادلة المقترحة لإيجاد حجم العينة المراد سحبها من مجتمع طبيعي تساوي إحدى العلاقات السابقة بعد حساب قيمة n ثم نطبق عليها صيغة ATI ولنأخذ الصيغة الأولى فهي تعطينا حجم العينة الأولي مثلها مثل المعادلات الأخرى ، أما حجم العينة النهائي المراد سحبه فسنحصل عليه وفق علاقة ATI .

إن علاقة ATI تعطينا نتائج أفضل من حجم العينة الأولي وتقدم معلومات عن المجتمع بشكل أكبر وأفضل لزيادة حجم العينة ، لكن يعيبها تكلفتها العالية في الفحص مقارنة مع العلاقات السابقة ويمكن أن نعتمد على الصيغ السابقة في حالة مراقبة الجودة إذا كانت نتائج الفحص تدميرية مثل فحص عمر المصابيح الكهربائية .

العلاقات السابقة تقف كلها عند حجم عينة ثابتة وهو حد الاشباع الثاني ، ويمكن تسمية حجم العينة النهائي بحجم العينة المعدل أو المصحح وفقاً لـ ATI .

ولو أردنا إيجاد العلاقة الرياضية بين حجم المجتمع ونسبة حجم العينة من الجدول (1) بعيداً عن التقريب لأعلى رقم صحيح فسنحصل على معادلة القوة التالية :

$$P_n = 23.05N^{(-0.67)} \qquad (11)$$

أما معامل التحديد فيساوي : $R^2 = 0.948$

تصلح عندما يكون حجم المجتمع 1000 مفردة وما فوق للسبب نفسه في المعادلة (8) .

أما لو أردنا إيجاد معادلة التمثيل بين حجم المجتمع ومتوسط اجمالي الفحص مباشرة دون الدخول بحجم العينة مباشرة لحصلنا على العلاقة الخطية التالية ومعامل تحديدها بالاعتماد على برنامج EXCEL:

$$ATI = 0.058N + 242.2$$
$$R^2 = 0.952 \qquad (12)$$
$$R = 0.976$$

وهي معادلة خطية فيها معامل الارتباط الخطي كبير جدا .

من خلال ما تقدم نلاحظ أيضاً أن معامل التحديد العائد لمعادلة متوسط إجمالي الفحص مع حجم المجتمع هو أكبر من معامل التحديد لمعادلة حجم العينة مع حجم المجتمع .

إن نسبة حجم العينة من حجم المجتمع عند حد الاشباع الأول تساوي تقريبا (1.3%) بينما تصبح هذه النسبة وفقا لـ ATI مساوية (6.2%) وهي بهذا تقدم معلومات أكثر ، لكن تحتاج لتكلفة أكبر وجهد أكبر .

يفضل استخدامها من الحجم 1000 مفردة وما فوق أيضاً .

سنبين في الجداول التالية نتائج تحليل التباين للانحدار للتأكد من صلاحية النموذج للعلاقة بين حجم المجتمع و متوسط إجمالي الفحص بالاعتماد على برنامج SPSS.

جدول رقم (5) يظهر معامل التحديد وخطأ التقدير

R	R Square	Adjusted R Square	Std. Error of the Estimate
.976	.952	.952	75.912

The independent variable is population size.

يبين الجدول رقم (5) أن قيمة معامل الارتباط بين حجم المجتمع ومتوسط إجمالي الفحص بلغت (0.976) ، وهي تدل على أن هناك علاقة ارتباط متينة جداً وطردية بينهما ، وهذا يدل على أن زيادة حجم المجتمع يؤدي إلى زيادة متوسط إجمالي الفحص .حيث أن حجم المجتمع في هذا النموذج يفسر ما مقداره (95.2%) من التباين الحاصل في زيادة متوسط إجمالي الفحص .

وبإعداد جدول تحليل ANOVA نحصل على الجدول رقم (6) التالي :

جدول رقم (6)ANOVA

	Sum of Squares	df	Mean Square	F	Sig.
Regression	1.059E7	1	1.059E7	1836.921	.000
Residual	530156.788	92	5762.574		
Total	1.112E7	93			

The independent variable is population size.

وبما أن مستوى الدلالة sig=0.000<α=0.05 فإننا نقول أن هناك أثر ذو دلالة معنوية لحجم المجتمع في متوسط إجمالي الفحص . أيضاً نلاحظ ثبات صلاحية النموذج استناداً إلى ارتفاع قيمة (F) المحسوبة والبالغة (1836.921) عن قيمتها الجدولية عند مستوى دلالة (0.05) ودرجات حرية (92,1) والبالغة (2.53) .

جدول رقم (7) ثوابت معادلة التمثيل واختبارها

	Unstandardized Coefficients		Standardized Coefficients	t	Sig.
	B	Std. Error	Beta		
population size	.059	.001	.976	42.859	.000
(Constant)	242.212	10.276		23.570	.000

نلاحظ من الجدول رقم (7) أن المعادلة الممثلة للعلاقة بين حجم المجتمع ومتوسط إجمالي الفحص تأخذ شكل معادلة خطية وهي نفس المعادلة رقم (12) التي حصلنا عليها من MS EXCEL ويعود الفرق للتقريب وتأخذ الشكل التالي :

$$ATI = 0.059N + 242.212 \qquad (12\text{-A})$$

كذلك نجد من الجدول أن قيمة sig المقابلة للثابت تساوي صفر بينما قيمة مستوى الدلالة $\alpha=0.05$ بالمقارنة نجد أن قيمة $Sig=0.000<\alpha=0.05$ هذا يعني أن قيمة الثابت معنوية.

الاستنتاجات والتوصيات:

الاستنتاجات:

1- إن النتائج كانت متطابقة بالنسبة للصيغ ذات الصلة بحجم المجتمع وحجم العينة عند توحيد شروطها، أي لا توجد فروق معنوية بين الصيغ المختلفة لحساب حجم العينة .

2- لم يزدد حجم العينة معنوياً مع زيادة حجم المجتمع عند حد الاشباع الأول ، ووجدنا أن حد الاشباع الأول المقابل لـ (30000) مفردة للمجتمع يقابل حجم للعينة يساوي 380 مفردة عند جميع الصيغ ، ما عدا صيغة بارتليت لحجم العينة المصحح أو النهائي .

3- لا توجد فروق معنوية بين حجمي العينة وفقا لحجم المجتمع عند حدي الاشباع ، إذ كان حد الاشباع الثاني لحجم المجتمع يساوي (1000000) مفردة يقابله حجم عينة يساوي 385 مفردة مهما كبر حجم المجتمع عن ذلك حتى ولو كان لا نهاية فهو سيبقى ثابتاً عند نفس القيمة لحجم العينة وهو 385 مفردة ولجميع الصيغ بما فيها علاقة بارتليت .

4- توجد فروق معنوية بين حجم العينة ومتوسط إجمالي الفحص وفقاً لحجم المجتمع عند حدي الاشباع ، وهذه الفروق تتضح من خلال حصولنا على قيمة متوسط إجمالي الفحص المقابل لحد الاشباع الأول يساوي ATI=1861 مفردة مقابل 380 مفردة لحجم العينة وفق الصيغ المستخدمة لحجم العينة ، بينما كانت قيمته عند حد الاشباع الثاني تساوي ATI = 50366 مفردة ولا يتوقف الأمر هنا بل مع كل زيادة لحجم المجتمع سيزداد ATI حتى لو تجاوز حجم المجتمع حد الاشباع الثاني .

5- حصلنا على نموذج رياضي للعلاقة بين حجم العينة وحجم المجتمع تم تمثيله بمعادلة لوغاريتمية وكان معامل تحديدها 0.927 وكانت قيمة الثابت معنوية ، علما أننا قربنا حجم العينة إلى أعلى رقم صحيح.

6- حصلنا على نموذج رياضي للعلاقة بين حجم المجتمع و نسبة حجم العينة تم تمثيله بمعادلة قوة وكان معامل تحديدها 0.948 دون تقريب لقيم النسبة .

7- حصلنا على نموذج رياضي للعلاقة بين حجم المجتمع ومتوسط اجمالي الفحص تم تمثيله بمعادلة خطية وكان معامل تحديدها 0.952 ومعامل الارتباط الخطي 0.976 مما يعكس العلاقة بين حجم المجتمع ومتوسط اجمالي الفحص بأنها طردية ومتينة جداً، وكانت قيمة الثابت معنوية.

8- وضعنا جدولا شاملا يعطينا حجم العينة المقابل لحجم المجتمع يمكن أن يكون في متناول الباحثين للاستفادة منه والاستغناء عن تطبيق الصيغ طالما هو يعتمد عليها بالأصل عند شروط معينة .

التوصيات:

1- بعد التأكد من تطابق النتائج عند توحيد الشروط لحجم العينة يجب الأخذ بعين الاعتبار لتطبيق صيغة واحدة يتم تداولها بين الباحثين وبخاصة طلاب الدراسات العليا إذا لم يكن الجدول بمتناول الباحثين .

2- الاستفادة من تطبيق النماذج الرياضية إن أمكن أيضاً في حساب حجم العينة وبخاصة عندما يكون حجم العينة مساوياً لـ 1000 مفردة وما فوق.

3- الاستفادة من تطبيق معادلة متوسط اجمالي الفحص وبخاصة في أبحاث مراقبة الجودة أو عندما يكون حجم المجتمع كبيرا .

4- توزيع جدول حساب حجم العينة على المستفيدين منه أو وضعه في متناولهم مثل طلاب الدراسات العليا وتوسيع عدد الجداول ليشمل جداول أخرى بشروط مختلفة عن الجدول الذي حصلنا عليه من البحث .

المراجع:

1- العلي ، إبراهيم ؛ عكروش ، محمد ، 2005 ، الإحصاء التطبيقي ، جامعة تشرين ، 175 .

2- العلي ، إبراهيم ؛ علي ، هارون ، 2004 ، الرياضيات العالية ، جامعة تشرين ، 231 .

3-القاضي ، دلال وأخرون ، 2005 ، الإحصاء للإداريين والاقتصاديين ، دار الحامد للنشر والتوزيع، الأردن- عمان ، 18 .

4-بسترفيلد، دال، 1995، الرقابة على الجودة، ترجمة ومراجعة سرور، سرور،المكتبة الأكاديمية، مصر القاهرة، 356 .

5- كبية، محمد؛ بدوي، ماهر، 2003 الإحصاء التطبيقي ، جامعة حلب ، 180 .

6- Bartlett, James E; Kotrlik,Joe W ., 2001 , Determining Appropriate Sample Size in Survey Research ,Information Technology,Learning, and performance Jornal , U. S.A.

7-Weimer Richard C.,1993-Statistics .(2nd ed) , W.M.C Brown publisher;U.S.A, 408.