وزارة التعليم العالي والبحث العلمي

جامعة زيان عاشور – الجلفة –

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

قسم العلوم التجارية

محاضرات في:

تقنيات المعاينة الإحصائية

موجهة لطلبة السنة الثانية ماستر

شعبة العلوم التجارية

تخصص: تسويق مصرفي

الاعداد:

الدكتور: طلال عمراني

أستاذ محاضر – أ–

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

جامعة الجلفة – الجزائر -

2022/2021

وزارة التعليم العالي والبحث العلمي

جامعة زيان عاشور – الجلفة –

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

قسم العلوم التجارية

محاضرات في:

تقنيات المعاينة الإحصائية

جامعة زيان عاشور بالجلفة
كلية العلوم الاقتصادية و التجارية و علوم التسيير

موجهة لطلبة السنة الثانية ماستر

شعبة العلوم التجارية

تخصص: تسويق مصرفي

الاعداد:

الدكتور: طلال عمراني

أستاذ محاضر – أ–

كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير

جامعة الجلفة – الجزائر -

2022/2021

الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة زيان عاشور بالجلفة

كلية العلوم الاقتصادية والتجارية وعلوم التسيير

قسم العلوم التجارية

جامعة زيان عاشور بالجلفة
مكتبة الكلية
كلية العلوم الاقتصادية و العلوم التجارية و علوم التسيير

مطبوعة محاضرات ودروس معمقة مقدمة لطلبة السنة الثانية ماستر

في

مقياس:

تقنيات المعاينة الإحصائية

إعداد الدكتور: طلال عمراني

تمهيد

لدراسة الخصائص المميزة لمجتمع ما، وعندما لا تتوفر لدينا البيانات اللازمة، يتطلب منا هذا ملاحظة وفحص عينة فقط من عناصر هذه المجموعة. كيفية جمع هذه البيانات تعد جوهر موضوع نظرية رياضية تسمى نظرية المعاينة (techniques d'échantillonnage)، هذه النظرية تسعى وبالشكل الأمثل لجمع البيانات على أساس معايير مختلفة وكذا الرد على بعض التساؤلات حول كيفية المضي قدما في عملية جمع البيانات بالمقارنة مع المعلومات المتاحة وجهد المعاينة المبذول.

تجدر الاشارة إلى اختلاف أسلوب المعاينة عن أسلوب الحصر الشامل (recensement)، إذ أن هذا الأخير يتم فيه جمع البيانات عن جميع مفردات المجتمع موضوع الدراسة، غير أنه عادة ما يكون من المستحيل الاتصال بجميع مفردات المجتمع من أجل عرض الاستبيان عليهم. فيتم اللجوء إلى أخذ عينة يمكن أن تؤدي إلى تحقيق تعميم النتائج على جميع مفردات المجتمع.

محتويات المطبوعة

- المحور الأول: مبادئ في الاحتمالات والاحصاء الرياضي
- المحور الثاني: جمع البيانات وطرق المعاينة
- المحور الثالث: العينة العشوائية البسيطة
- المحور الرابع: العينة العشوائية الطبقية
- المحور الخامس: العينة العشوائية المنتظمة
- المحور السادس: العينة العشوائية العنقودية من مرحلة واحدة
- المحور السابع: العينة العشوائية العنقودية من مرحلتين

المحور الأول: مبادئ في الاحتمالات والاحصاء الرياضي

سنقوم في هذا المحور بتقديم مختلف المفاهيم والتقنيات الاحصائية الواجب معرفتها للتمكن من فهم موضوع تقنيات المعاينة. إذ أنّ تطور أساليب وطرق المعاينة يعود بالدرجة الأولى إلى تطور نظرية الاحتمالات التي تقوم على دراسة وتحليل مختلف الظواهر العشوائية.

1. المتغير العشوائي

يعرف على أنه مقدار يمكن له أخذ مجموعة من القيم باحتمالات معينة، وهو معرف على مجموعة القيم الممكنة لتجربة عشوائية (رمي قطعة نقد، رمي النرد، عائد سهم معين، ...).،حيث يكون مجموع هذه الإحتمالات يساوي الواحد (1).

والمتغير العشوائي يمكن أن يكون إما منفصلا وإما متصلا، منفصلا إذا كان يأخذ مجموعة من النتائج يمكن إحصاءها، كما قد تكون غير قابلة للإحصاء ولكن في مجموعة الأعداد الطبيعية $x \in N$، مثل قانون بواسون، ويكون المتغير العشوائي متصلا إذا كانت قيمه الممكنة معرفة على مجموعة الأعداد الحقيقية $x \in R$. وبالنسبة لقانون الاحتمال، فإننا نميز بين دالة الاحتمالات $p(x)$ بالنسبة للمتغير المنفصل، ودالة الكثافة $f(x)$ بالنسبة للمتصل.

2. الأمل الرياضي (المتوسط) والتباين

إذا كان X متغير عشوائي منفصل يأخذ القيم x_i بالاحتمالات $p(x_i)$، فإنه من أجل كل دالة g لدينا:

$$E\big(g(X)\big) = \sum_i g(x_i) \cdot p(x_i)$$

وتصبح العلاقة السابقة كما يلي:

$$E\big(g(X)\big) = \int_{df} g(x).\, f(x).\, dx$$

إذا كان X متغير عشوائي متصل يأخذ قيمه في مجال تعريف df.

كما يمكن إستخلاص العلاقات الآتية:

- $E(X) = \sum_i x_i \,.\, p(x_i) = \mu_X$
- $E(X^n) = \sum_i x_i^n \,.\, p(x_i)$
- $E(a) = a$ حيث a عبارة عن ثابت.
- $E(aX) = a.E(X)$

- $E(aX+b)=a.E(X)+b$ حيث a,b عبارة عن ثوابت.
- $E(X+Y)=E(X)+E(Y)$
- $E(\sum_{i=1}^{n}X_i)=\sum_{i=1}^{n}E(X)=n.E(X)$ حيث تكون قيم X_i مستقلة ولها نفس دالة الإحتمال أو الكثافة.

ويعطى تباين X بالعلاقة:

$$V(X)=E((X-E(X))^2)=E(X^2)-E(X)^2=\delta_X^2$$

كما يمكن إستخلاص العلاقات الآتية:

- $V(b)=0$
- $V(aX)=a^2.V(X)$
- $V(aX+b)=V(aX)+V(b)=a^2.V(X)$
- $V(X+Y)=V(X)+V(Y)+2.COV(X,Y)$ حيث:
  $COV(X,Y)=E\big((X-E(X)\big).(Y-E(Y)))$
- $V(X+Y)=V(X)+V(Y)$ إذا كان X, Y متغيرين مستقلين.
- $V(\sum_{i=1}^{n}X_i)=\sum_{i=1}^{n}V(X)=n.V(X)$ حيث تكون قيم X_i مستقلة ولها نفس دالة الإحتمال أو الكثافة.
- ويعطى الانحراف المعياري بالعلاقة:

$$\delta=\sqrt{V(X)}$$

مثال 1-1:

لنفترض أن المتغير العشوائي X يمثل مختلف النتائج الممكن الحصول عليها من رمي نرد متزن.

المطلوب: أحسب الأمل الرياضي وتباين المتغير X.

الحل:

النتائج الممكن الحصول عليها من رمي النرد هي: $\{1,2,3,4,5,6\}$ باحتمال يساوي $\frac{1}{6}$ بالنسبة لكل نتيجة ممكنة، وعليه:

$$E(X)=1.\frac{1}{6}+2.\frac{1}{6}+3.\frac{1}{6}+4.\frac{1}{6}+5.\frac{1}{6}+6.\frac{1}{6}=\frac{7}{2}$$

$$E(X^2)=1^2.\frac{1}{6}+2^2.\frac{1}{6}+3^2.\frac{1}{6}+4^2.\frac{1}{6}+5^2.\frac{1}{6}+6^2.\frac{1}{6}=\frac{91}{6}$$

$$V(X) = \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12}$$

3. قانون برنولي B(p):

يمثل قانون احتمال المتغير العشوائي المنفصل x الذي يعطي نتيجتين فقط (مثل: رمي قطعة النقد،...)، ويعطى قانون الاحتمالات بالشكل:

$$P(X=x) = \begin{cases} p & si\ x = 1 \\ q = 1 - p & si\ x = 0 \end{cases}$$

وعبارتي الأمل الرياضي والتباين كما يلي:

$$E(X) = p$$

$$V(X) = p.(1-p) = p.q \ \rightarrow \ \delta = \sqrt{p.q}$$

4. قانون ثنائي الحد لبينوميال B(n, p):

يمثل قانون احتمال المتغير العشوائي المنفصل x الذي يعطي نتيجة محددة (ظهور وجه قطعة نقد) خلال n سحب، مستقل وبالإرجاع، في صندوق يحوي N كرة فيها نسبة معينة p% من الكريات البيضاء (مثل: رمي قطعة النقد n مرة،...)، ويعطى قانون الاحتمالات بالشكل:

$$P(X = k) = C_n^k . p^k . (1-p)^{n-k} \ avec\ k \in \{0,1,2,\dots,k\} / \ k \leq n$$

وعبارتي الأمل الرياضي والتباين كما يلي:

$$E(X) = n.p$$

$$V(X) = n.p.(1-p) = n.p.q \rightarrow \ \delta = \sqrt{n.p.q}$$

$$C_n^k = \frac{n!}{k!.(n-k)!}$$

في الأخير، تجدر الإشارة إلى أن قانون بينوميال B(n, p) يمثل تكرار لقانون برنولي n مرة بشكل مستقل وبنفس الاحتمال p.

5. قانون بواسون $P(\lambda)$

يعتبر من أهم التوزيعات الإحتمالية المنفصلة، يستخدم في المجالات التي يكون هدفها إيجاد توزيعات عدد المرات التي تحدث فيها ظاهرة عشوائية خلال فترة معينة (الظواهر النادرة ذات

التكرار المتنازل)، كعدد الأخطاء المطبعية في صفحة من كتاب محدد، ونكتب في هذه الحالة : $x \sim P(\lambda)$. أي أن X يتبع توزيع بواسون المعرف على (λ)، وهي معلمة التوزيع، قيمتها موجبة وتمثل المتوسط الحسابي، ويعطى قانون الاحتمالات بالشكل:

$$P(X = k) = e^{-\lambda}.\frac{\lambda^k}{k!} \text{avec } \lambda > 0 \text{ et } k \in N.$$

وعبارتي الأمل الرياضي والتباين كما يلي:

$E(X) = \lambda$;

$V(X) = \lambda$.

6. تقريب توزيع ثنائي الحد بتوزيع بواسون

يمكن إستخدام توزيع بواسون كتقريب لتوزيع ثنائي الحد عندما يكون عدد التجارب كبيرا جدا $(n \geq 30)$ والإحتمال p صغيرا جدا$(p \leq 0{,}1)$ ، بحيث يكون $(n.p.q \leq 10)$. ويهدف هذا التقريب لتسهيل عملية حساب الإحتمالات، فيصبح X يتبع توزيع بواسون بمعلمة $(\lambda = n.p)$.

مثال 1-2:

إذا كانت 5% من الوحدات المنتجة بواسطة أحد الخطوط الإنتاجية لشركة ما غير سليمة.

المطلوب:

- ماهو إحتمال الحصول على 05 وحدات منتجة غير سليمة من عينة من 120 وحدة ؟ (بإستخدام توزيع ذي الحدين وكذا بواسون كتقريب لذي الحدين)

الحل:

n=120, k=5, p=0,05, q=1-0,05=0,95.

بإستخدام علاقة التوزيع الإحتمالي لذي الحدين:

$$P(X = 5) = C_{120}^{5}.(0{,}05)^5.(1 - 0{,}05)^{120-5}$$

$$= \frac{120!}{5!.(120-5)!}.(0{,}05)^5.(0{,}95)^{115}$$

$$= 0{,}1633$$

بإستخدام علاقة التوزيع الإحتمالي لبواسون:

$(n = 120 \geq 30)$ و $(p = 0{,}05 \leq 0{,}1)$ و $(n.p.q = 5{,}7 \leq 10)$

حيث: $\lambda = n.p = 120.(0{,}05) = 6$

$$P(X = 5) = e^{-6}.\frac{6^5}{5!} = 0{,}1606$$

بتطبيق القانونين نلاحظ أن النتيجتين متساويتين بالتقريب.

7. تقريب التوزيع فوق الهندسي بتوزيع ثنائي الحد

نقول أن X يتبع التوزيع فوق الهندسي $x \sim H(N, n, t)$ ، حيث:

N : حجم المجتمع، n : حجم العينة، t : عدد وحدات الظاهرة محل الدراسة في المجتمع.

ويعطى قانون الاحتمالات بالشكل:

$$P(X = k) = \frac{C_{N-t}^{n-k}.C_t^k}{C_N^n}$$

وتمثل k عدد وحدات الظاهرة محل الدراسة في العينة.

إذا كانت التجربة بمتغير عشوائي فيه شروط التوزيع فوق الهندسي (سحب بدون إرجاع لعينة من مجتمع وكان حجم العينة صغيرا جدا بالنسبة للمجتمع $\frac{n}{N} < 0{,}1$)، فإنه يمكن تقريب التوزيع فوق الهندسي إلى التوزيع ثنائي الحد، حيث يصبح إحتمال النجاح الموافق للظاهرة محل الدراسة $p = \frac{t}{N}$، وفي هذه الحالة نحصل على نتيجتين متساويتين بالتقريب عند تطبيق القانونين.

مثال 1-3:

أحسب إحتمال إختيار 3 وحدات سليمة من عينة مكونة من 5 وحدات أخذت عشوائيا وبدون إرجاع من مجموعة مكونة من 500 وحدة منها 250 سليمة.

الحل:

N=500, t=250, n=5, k=3.

بإستخدام علاقة التوزيع الإحتمالي فوق الهندسي:

$$P(X=3)=\frac{C_{500-250}^{5-3}C_{250}^{3}}{C_{500}^{5}}=\frac{\frac{250!}{2!.248!}\frac{250!}{3!.247!}}{\frac{500!}{5!.495!}}=0{,}3137$$

بإستخدام علاقة التوزيع الإحتمالي ثنائي الحد: $(\frac{n}{N}=\frac{5}{500}=0{,}01<0{,}1)$

$$p=\frac{t}{N}=\frac{250}{500}=0{,}5$$

$$P(X=k)=C_{n}^{k}.p^{k}.(1-p)^{n-k}$$

$$P(X=3)=C_{5}^{3}.(0{,}5)^{3}.(0{,}5)^{5-3}=0{,}3125$$

8. القانون الأسي $\mathcal{E}(\lambda)$

يعتبر من أهم التوزيعات الإحتمالية المتصلة، يستخدم في المجالات التي يكون هدفها إيجاد توزيعات مدة بقاء الظاهرة العشوائية محل الدراسة، كمدة إنتظار شخص ما، مدة صلاحية منتوج معين، ...، ونكتب في هذه الحالة : $x\sim\mathcal{E}(\lambda)$. أي أن X يتبع التوزيع الأسي المعرف على (λ)، هذه الأخيرة تأخذ قيما موجبة وتسمى بمعلمة أو وسيط التوزيع، وتعطى دالة كثافته بالشكل:

$$f(x)=\lambda.e^{-\lambda x}\ \text{avec}\ x\geq 0$$

وعبارتي الأمل الرياضي والتباين كما يلي:

$$E(X)=\frac{1}{\lambda}\ ;$$

$$V(X)=\frac{1}{\lambda^{2}}$$

مثال 1-4:

مدة صلاحية مكون إلكتروني مقدرة بالساعات، تتبع التوزيع الأسي بوسيط قدره 0,005.

- ماهي مدة الصلاحية المتوسطة لهذا المكون الإلكتروني؟
- أكتب دالة الكثافة لهذا المتغير؟
- أحسب إحتمال أن تكون مدة صلاحية مكون إلكتروني مأخوذ عشوائيا أقل من 100 ساعة؟
- ماهو إحتمال أن يبقى هذا المكون صالحا للإستعمال لأكثر من 250 ساعة؟

الحل:

1/ مدة الصلاحية المتوسطة للمكون الإلكتروني هي:

$$E(X) = \frac{1}{\lambda} = \frac{1}{0{,}005} = 200$$

2/ دالة الكثافة:

$$\mathrm{f(x)} = 0{,}005.\,\mathrm{e}^{-0{,}005.x} \text{ avec } \mathrm{x} \geq 0$$

3/ إحتمال أن تقل مدة صلاحية المكون الإلكتروني عن 100 ساعة:

$$\mathrm{P(X \leq 100)} = \int_0^{100} 0{,}005.\,\mathrm{e}^{-0{,}005.x}.\,dx = [-\mathrm{e}^{-0{,}005.x}]_0^{100}$$
$$= -\mathrm{e}^{-0{,}005.(100)} + \mathrm{e}^{-0{,}005.(0)} = 1 - \mathrm{e}^{-0{,}5} = 0{,}39.$$

هناك إحتمال 39% أن تقل مدة صلاحية المكون الإلكتروني عن 100 ساعة.

4/ إحتمال أن تزيد مدة صلاحية المكون الإلكتروني عن 250 ساعة:

$$\mathrm{P(X \geq 250)} = \int_{250}^{+\infty} 0{,}005.\,\mathrm{e}^{-0{,}005.x}.\,dx = [-\mathrm{e}^{-0{,}005.x}]_{250}^{+\infty}$$
$$= -\mathrm{e}^{-0{,}005.(+\infty)} + \mathrm{e}^{-0{,}005.(250)} = -\mathrm{e}^{-\infty} + \mathrm{e}^{-1{,}25} = \mathrm{e}^{-1{,}25}$$
$$= 0{,}29.$$

هناك إحتمال قدره 29% أن تزيد مدة صلاحية المكون الإلكتروني عن 250 ساعة.

9. التوزيع الطبيعي $N(\mu, \delta^2)$ والتوزيع الطبيعي المعياري أو القياسي المنتظم $N(0, 1)$

يعتبر هذا التوزيع من أهم التوزيعات الإحتمالية المستعملة على نطاق واسع، ويتعلق بالمتغير العشوائي المتصل x، المعرف على المجال $]-\infty, +\infty[$، حيث تعطى دالة الكثافة بالشكل:[1]

$$f(x) = \frac{1}{\delta.\sqrt{2\pi}}.e^{-\frac{1}{2}.\left(\frac{x-\mu}{\delta}\right)^2}$$

حيث: $E(x) = \mu$ و$V(x) = \delta^2$ يمثلان على التوالي متوسط وتباين المتغير العشوائي x، وعليه نكتب $x \sim N(\mu, \delta^2)$.

الدالة $f(x)$ مستمرة وموجبة على مجال تعريفها، كما أن مجموع الإحتمالات الموافقة لجميع قيم المتغير العشوائي x يساوي الواحد. $(\int_{-\infty}^{+\infty} f(x).dx = 1)$.

نشير أيضا إلى أن دالة التوزيع الاحتمالي الطبيعي للمتغير العشوائي تعطى بالعلاقة:

$$F_X(x) = P(X \leq x)$$

وهي دالة متزايدة تأخذ قيما بين 0 و1.

وتمثيل بياني على شكل منحنى جرس (courbe en cloche) متناظر بالنسبة للمحور $x = \mu$ كما هو موضح في الشكل 1-1 أسفله.

إذا كان لدينا المتغير العشوائي $x \sim N(\mu, \delta^2)$، فإنه يمكن التحول إلى التوزيع الطبيعي المعياري، حيث يتم تحويل المتغير x إلى متغير معياري z حيث:

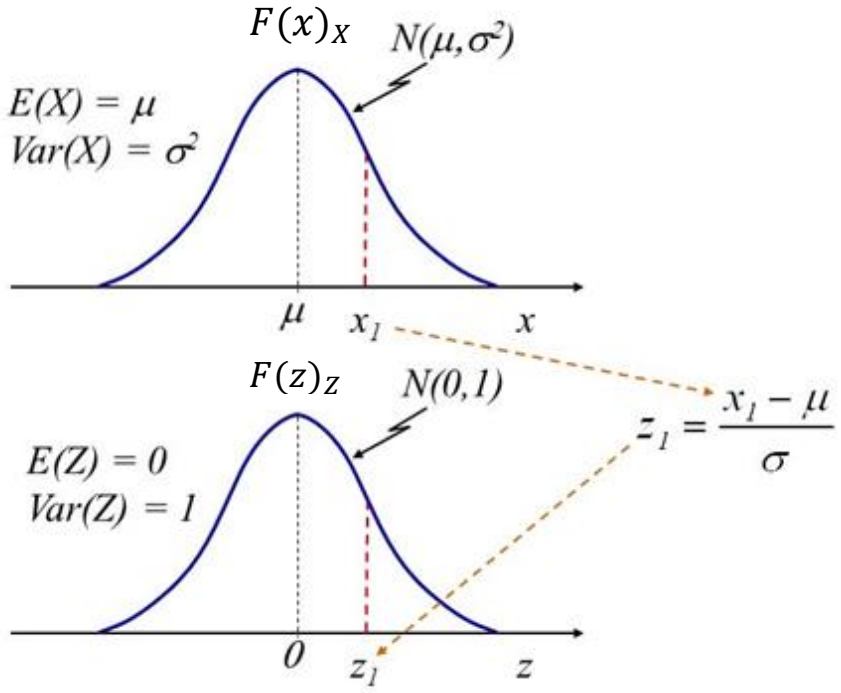
$$z = \frac{x - \mu}{\delta} \sim N(0, 1)$$

وتصبح قيمتي المتوسط μ معدومة والتباين δ^2 تساوي الواحد.

والتمثيل البياني على شكل منحنى جرس (courbe en cloche) متناظر بالنسبة للمحور $x = 0$ كما يلي:

[1] جابر أحمد بسيوني، "الإحصاء العام"، دار الوفاء لدنيا الطباعة والنشر، الاسكندرية 2013، ص226.

الشكل رقم 1-1: المنحنى البياني لدالتي التوزيع الاحتمالي الطبيعي والطبيعي المعياري



المصدر: من إعداد الأستاذ

وللحصول على مختلف قيم الاحتمالات لدالة $F(z)_Z$ نستعمل جدول التوزيع الطبيعي $\mathrm{N}(0,1)$ (أنظر الجدول في الملحق)، وبإستخدام الخصائص التالية:

1/ $\mathrm{P}(\mathrm{Z} \geq 0) = \mathrm{P}(\mathrm{Z} \leq 0) = 0{,}5$;

2/ $\mathrm{P}(\mathrm{Z} \geq \mathrm{z}) = 1 - \mathrm{P}(\mathrm{Z} \leq \mathrm{z})$;

3/ $\mathrm{P}(\mathrm{Z} \leq -\mathrm{z}) = 1 - \mathrm{P}(\mathrm{Z} \leq +\mathrm{z})$;

4/ $\mathrm{P}(-\mathrm{z}_1 \leq \mathrm{Z} \leq +\mathrm{z}_2) = \mathrm{P}(\mathrm{Z} \leq +\mathrm{z}_2) - \mathrm{P}(\mathrm{Z} \leq -\mathrm{z}_1)$;

5/ $\mathrm{P}(-\mathrm{z} \leq \mathrm{Z} \leq +\mathrm{z}) = 2.\mathrm{P}(\mathrm{Z} \leq +\mathrm{z}) - 1$.

ملاحظة: عندما تكون قيمة n صغيرة $(\mathrm{n} < 30)$، فإن مختلف قيم الاحتمالات لدالة التوزيع الاحتمالي تستخرج من جدول توزيع ستودينت بدرجة حرية n-1 $(\mathrm{t}^{\alpha}_{\mathrm{n}-1})$ (أنظر الجدول في الملحق).

مثال 1-5:

بفرض أن دخل الأسر يتبع التوزيع الطبيعي مع $\mu = 16.000$ و$\delta = 2.000$، ماهو إحتمال أن يكون دخل أسرة أختيرت عشوائيا بين 15.000 دج و18.000 دج ؟

الحل:

$x \sim N(16.000, 2.000^2) \rightarrow z = \frac{x-\mu}{\delta} \sim N(0,1)$;

$P(15.000 \leq x \leq 18.000) = ?$

لإستخدام المتغير الطبيعي المعياري z نقوم بالتغييرات الآتية:

$$P(15.000 \leq x \leq 18.000) = P\left(\frac{15.000 - 16.000}{2.000} \leq \frac{x-\mu}{\delta} \leq \frac{18.000 - 16.000}{2.000}\right)$$

$$= P(-0.5 \leq z \leq 1)$$

$$= P(z \leq 1) - P(z \leq -0,5) = P(z \leq 1) - 1 + P(z \leq 0,5)$$

بإستعمال جدول التوزيع الطبيعي المعياري $N(0,1)$ (أنظر الجدول في الملحق)، نستخرج القيم الموافقة فنجد:

$$= 0,8413 - 1 + 0,6915 = 0,5328 = 53,28\%$$

لدينا إحتمال 53,28% أن يتراوح دخل الأسرة المختارة بين 15.000دج و18.000دج.

10. قانون الأعداد الكبيرة (loi des grands nombres)

إذا كانت $(X_1, X_2, \dots, X_n)$ مجموعة متغيرات مستقلة، متماثلة وبنفس التوزيع (i.i.d) حسب قانون لا على التعيين بمتوسط μ فإن:

$$\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i \underset{n\to\infty}{\overset{p.s}{\rightarrow}} \mu$$

بشكل آخر، متوسط المتغير العشوائي لعينة عشوائية بسيطة يؤول إلى متوسط المجتمع، عندما يؤول حجم العينة إلى مالا نهاية. على سبيل المثال، إذا كان نرمي قطعة نقد سوية بشكل متعدد، فإن احتمال ظهور الوجهين يؤول إلى النصف $\left(\frac{1}{2}\right)$.

11. نظرية النهاية المركزية (théorème central limite)

إذا كانت $(X_1, X_2, \dots, X_n)$ مجموعة متغيرات مستقلة، متماثلة وبنفس التوزيع (i.i.d) حسب قانون لا على التعيين بمتوسط μ وتباين δ^2 فإن:

$$\frac{\overline{X}_n - \mu}{\frac{\delta}{\sqrt{n}}} \underset{n\to\infty}{\rightarrow} N(0,1)$$

من الناحية العملية فإن n تؤول لما نهاية أو تكون كبيرة جدا عندما تكون $n > 30$.[2]

12. مجال الثقة (intervalle de confiance - IC)

عادة مايكون من الصعب تقدير القيمة الحقيقية لمعلمة ما (لتكن θ)، ولهذا نقوم بإيجاد مجال ثقة لهذه المعلمة عوض القيمة الفعلية.

ومجال الثقة لهذه المعلمة θ عند مستوى الثقة $(1-\alpha)$ هو كل مجال IC حيث يكون:

$$P(\theta \in IC) = 1 - \alpha$$

حيث $\alpha \in [0,1]$ ويمثل مستوى المعنوية.

نشير هنا إلى أنه كلما زادت قيمة α كلما قل مدى مجال الثقة.

إحتمال أن تكون القيمة الفعلية θ خارج مجال الثقة هو α، وهذه القيمة الأخيرة يمكن أن توزع بين طرفي المجال بقيم α_1 و α_2 حيث: $\alpha = \alpha_1 + \alpha_1$.

- مجال الثقة يكون ثنائي الحد عندما $\alpha_1 \neq 0$ و $\alpha_2 \neq 0$ ، ويكون ثنائي الحد متناظر لما $\alpha_2 = \alpha_1 = \frac{\alpha}{2}$ وإلا غير متناظر إذا كان $\alpha_2 \neq \alpha_1$.
- إذا كان مجال الثقة محدود بحد أعلى فإن $\alpha_1 = 0$ و $\alpha_2 = \alpha$ ويكون المجال في هذه الحالة بالشكل: $IC \in]-\infty, c]$.
- إذا كان مجال الثقة محدود بحد أدنى فإن $\alpha_1 = \alpha$ و $\alpha_2 = 0$ ويكون المجال في هذه الحالة بالشكل: $IC \in [c, +\infty[$.

1.12. مجال الثقة بالنسبة للأمل الرياضي μ

ليكن لدينا $X \sim N(\mu, \delta^2)$ و$(X_1, X_2, \dots, X_n)$ مجموعة متغيرات مستقلة، متماثلة وبنفس التوزيع الاحتمالي.

[2] Régis BOURBONNAIS et Michel TERRAZA, "Analyse des séries temporelles en économie", Édition PUF, Paris 1998, Page 218.

الحالة الأولى: قيمة التباين δ^2 معلومة

مع العلم أن: $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$ وحسب نظرية النهاية المركزية (TCL) فإن:

$$Z = \frac{\overline{X}_n - \mu}{\frac{\delta}{\sqrt{n}}} \underset{n\to\infty}{\rightarrow} N(0,1)$$

لدينا: $P\left(-z_{1-\frac{\alpha}{2}} \leq \frac{\overline{X}_n - \mu}{\frac{\delta}{\sqrt{n}}} \leq +z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$

وتستخرج قيمة $z_{1-\frac{\alpha}{2}}$ من جدول التوزيع الطبيعي $N(0,1)$ (أنظر الجدول في الملحق).[3]

وعليه مجال الثقة ثنائي الحد المتناظر بالنسبة ل μ يكون:

$$\boxed{\overline{X}_n - z_{1-\frac{\alpha}{2}}.\frac{\delta}{\sqrt{n}} \leq \mu \leq \overline{X}_n + z_{1-\frac{\alpha}{2}}.\frac{\delta}{\sqrt{n}}}$$

الحالة الثانية: قيمة التباين δ^2 غير معلومة

في هذه الحالة نستعمل المقدرة غير المتحيزة $\hat{S}^2$ وهي: $\hat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

وعليه يصبح:

$$T = \frac{\overline{X}_n - \mu}{\frac{\hat{S}_n}{\sqrt{n}}} \sim t_{n-1}^{\frac{\alpha}{2}}$$

أو

$$T = \frac{\overline{X}_n - \mu}{\frac{S_n}{\sqrt{n-1}}} \sim t_{n-1}^{\frac{\alpha}{2}}$$

[3] Sandrine LARDIC et Valérie MIGNON, "Économétrie des séries temporelles macroéconomiques et financières", Édition ECONOMICA, Paris 2002, Page 361.

($t_{n-1}^{\frac{\alpha}{2}}$: التوزيع الاحتمالي لستيودنت بدرجة حرية n-1 ومستوى ثقة $1-\alpha$ (أنظر الجدول في الملحق))[4].

لدينا: $P\left(-t_{n-1}^{\frac{\alpha}{2}} \leq \frac{\overline{X}_n-\mu}{\frac{\hat{S}_n}{\sqrt{n}}} \leq +t_{n-1}^{\frac{\alpha}{2}}\right) = 1-\alpha$

وعليه مجال الثقة ثنائي الحد المتناظر بالنسبة ل μ يكون:

$$\overline{X}_n - t_{n-1}^{\frac{\alpha}{2}}.\frac{\hat{S}_n}{\sqrt{n}} \leq \mu \leq \overline{X}_n + t_{n-1}^{\frac{\alpha}{2}}.\frac{\hat{S}_n}{\sqrt{n}}$$

كما يمكن كتابته بالعبارة:

$$\overline{X}_n - t_{n-1}^{\frac{\alpha}{2}}.\frac{S_n}{\sqrt{n-1}} \leq \mu \leq \overline{X}_n + t_{n-1}^{\frac{\alpha}{2}}.\frac{S_n}{\sqrt{n-1}}$$

ملاحظة: عندما تكون قيمة n كبيرة، فإن توزيع ستيودنت يؤول إلى التوزيع الطبيعي $N(0,1)$، ويصبح مجال ثقة μ هو نفس مجال الثقة في الحالة الأولى.

2.12. مجال الثقة بالنسبة للتباين δ^2

ليكن لدينا $X \sim N(\mu, \delta^2)$ و$(X_1, X_2, \dots, X_n)$ مجموعة متغيرات مستقلة، متماثلة وبنفس التوزيع الاحتمالي.

الحالة الأولى: قيمة الأمل الرياضي μ معلومة

ليكن: $S_n^{*2} = \frac{1}{n}\sum_{i=1}^{n}(X_i-\mu)^2$

لدينا: $\frac{S_n^{*2}}{\frac{\delta^2}{n}} \sim \chi_n^2$ (التوزيع الاحتمالي كي دو بدرجة حرية n (أنظر الجدول في الملحق)).[5]

لدينا: $P\left(\chi_{\frac{\alpha_1}{2}}^2 \leq \frac{S_n^{*2}}{\frac{\delta^2}{n}} \leq \chi_{1-\frac{\alpha_2}{2}}^2\right) = 1-\alpha$

[4] Idem, Page 362.
[5] Sandrine LARDIC et Valérie MIGNON, Op cit, Page 363.

وعليه مجال الثقة ثنائي الحد بالنسبة ل δ^2 يكون (هنا المجال غير المتناظر لأن توزيع كي دو غير متناظر):

$$\frac{S_n^{*2}}{\frac{\chi^2_{1-\frac{\alpha_2}{2}}}{n}} \leq \delta^2 \leq \frac{S_n^{*2}}{\frac{\chi^2_{\frac{\alpha_1}{2}}}{n}}$$

الحالة الثانية: قيمة الأمل الرياضي μ غير معلومة

ليكن: $\hat{S}_n^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$

لدينا: $\frac{\hat{S}_n^2}{\frac{\delta^2}{n-1}} \sim \chi^2_{n-1}$ (التوزيع الاحتمالي كي دو بدرجة حرية n-1)

لدينا: $P\left(\chi^2_{\frac{\alpha_1}{2}} \leq \frac{\hat{S}_n^2}{\frac{\delta^2}{n-1}} \leq \chi^2_{1-\frac{\alpha_2}{2}}\right) = 1 - \alpha$

وعليه مجال الثقة ثنائي الحد بالنسبة ل δ^2 يكون:

$$\frac{\hat{S}_n^2}{\frac{\chi^2_{1-\frac{\alpha_2}{2}}}{n-1}} \leq \delta^2 \leq \frac{\hat{S}_n^2}{\frac{\chi^2_{\frac{\alpha_1}{2}}}{n-1}}$$

13. إختبار الفرضيات (tests d'hypothèses)

المبدأ الذي يرتكز عليه اختبار الفرضيات هو طرح فرضية العمل والتنبؤ بآثارها على المجتمع أو العينة. ثم مقارنة هذه التنبؤات مع الملاحظات واتخاذ قرار قبول أو رفض هذه الفرضية بناءا على أسس موضوعية.[6]

1.13. فرضية العدم والفرضية البديلة

فرضية العمل وتسمى فرضية العدم H_0 وهي دوما الفرضية المراد اختبارها. يتم في المقابل استخدام الرمز H_1 للدلالة على الفرض المغاير للفرض العدمي، ويكون اختبار الفرضيتين السابقتين على الشكل:[7]

[6] David R. ANDERSON, Jeffrey D. CAMM, James J. COCHRAN, Dennis J. SWEENEY, Thomas A. WILLIAMS, " Statistiques pour l'économie et la gestion", Édition DE BOECK, Paris 2015, P498.

[7] ثائر فيصل شاهر، "اختبار الفرضيات الاحصائية"، دار الحامد للنشر والتوزيع، عمان 2013، ص67.

$$\begin{cases} H_0 : \text{فرضية العدم (الفرضية التي سيتم إختبارها)} \\ H_1 : \text{الفرضية البديلة} \end{cases}$$

تتعلق عملية اختبار الفرضيات عادة بالقيمة الفعلية لمعلمة محددة كمتوسط متغير عشوائي μ، تباين δ^2 أو نسبة حدوث حدث معين P. والاختبار يكون في شكل إحدى هذه الحالات الثلاث:

$$\begin{cases} H_0 : \text{أن تكون قيمة معلمة المجتمع لا تختلف عن قيمة محددة} \\ H_1 : \text{أن تكون قيمة معلمة المجتمع مختلفة عن قيمة محددة} \end{cases}$$

$$\begin{cases} H_0 : \text{أن تكون قيمة معلمة المجتمع لا تزيد عن قيمة محددة} \\ H_1 : \text{أن تكون قيمة معلمة المجتمع أكثر من قيمة محددة} \end{cases}$$

$$\begin{cases} H_0 : \text{أن تكون قيمة معلمة المجتمع لا تقل عن قيمة محددة} \\ H_1 : \text{أن تكون قيمة معلمة المجتمع أقل من قيمة محددة} \end{cases}$$

2.13. الإختبار أحادي الحد والاختبار ثنائي الحد

ويسمى كذلك الاختبار من طرف واحد والاختبار من طرفين، فنلجأ للاختبار من طرفين عندما يكون إختبار الفرضيات من الشكل:

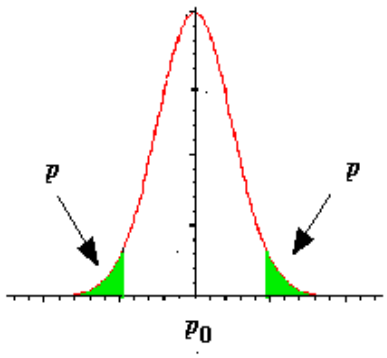
$$\begin{cases} H_0 : \text{أن تكون قيمة معلمة المجتمع لا تختلف عن قيمة محددة} \\ H_1 : \text{أن تكون قيمة معلمة المجتمع مختلفة عن قيمة محددة} \end{cases}$$

فعلى سبيل المثال، لاختبار الفرض العدمي H_0 القائل بأن نسبة المدخنين في مجتمع الطلبة P لا تختلف عن نسبة المدخنين في المجتمع ككل P_0، يكون الاختبار كالآتي:

$$\begin{cases} H_0 : P = P_0 \\ H_1 : P \neq P_0 \end{cases}$$

يكون الاختبار في هذه الحالة من طرفين لأن النسبة P في الحقيقة قد تكون أكبر أو صغر من P_0 (أنظر الشكل رقم 02).

الشكل رقم 1-2: الاختبار ثنائي الطرفين



المصدر: من إعداد الأستاذ

في المقابل، يكون الاختبار من طرف واحد وذلك إما على اليمين وإما على اليسار إذا كان إختبار الفرضيات من الشكلين:

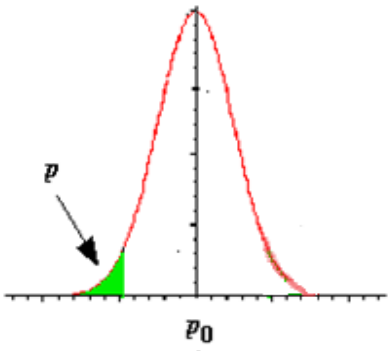
$$\begin{cases} H_0 : \text{أن تكون قيمة معلمة المجتمع لا تزيد عن قيمة محددة} \\ H_1 : \text{أن تكون قيمة معلمة المجتمع أكثر من قيمة محددة} \end{cases}$$

$$\begin{cases} H_0 : \text{أن تكون قيمة معلمة المجتمع لا تقل عن قيمة محددة} \\ H_1 : \text{أن تكون قيمة معلمة المجتمع أقل من قيمة محددة} \end{cases}$$

فإذا أردنا مثلا إختبار الفرض العدمي H_0 القائل بأن نسبة المدخنين في مجتمع الطلبة P لا تقل عن نسبة المدخنين في المجتمع ككل P_0، يكون الاختبار من طرف واحد على اليمين (أنظر الشكل رقم 03) كالآتي:

$$\begin{cases} H_0 : P \geq P_0 \\ H_1 : P < P_0 \end{cases}$$

الشكل رقم 1-3: الاختبار أحادي الطرف على اليمين

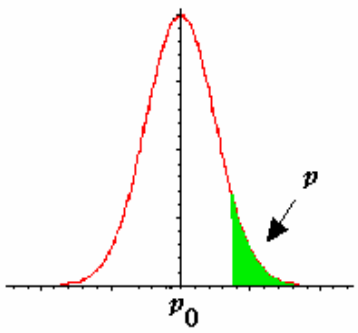


المصدر: من إعداد الأستاذ

أما إذا أردنا إختبار الفرض العدمي H_0 القائل بأن نسبة المدخنين في مجتمع الطلبة P لا تزيد عن نسبة المدخنين في المجتمع ككل P_0، يكون الاختبار من طرف واحد على اليسار (أنظر الشكل رقم 04) كالآتي:

$$\begin{cases} H_0 : P \leq P_0 \\ H_1 : P > P_0 \end{cases}$$

الشكل رقم 1-4: الاختبار أحادي الطرف على اليسار



المصدر: من إعداد الأستاذ

3.13. الدالة الاختبارية وقرار القبول أو الرفض

فيما يلي الدالة الاختبارية لكل من μ وP.

1.3.13. الدالة الاختبارية لمتوسط المجتمع μ

للقيام باختبار الفرضيات عند مستوى ثقة (1-α)، نقوم باحتساب الدالة الاختبارية $Z = \frac{\bar{x}-\mu}{\hat{\delta}_{\bar{x}}}$ ونرى إن كانت تنتمي لمنطقة القبول (zone d'acceptation) أو الرفض:

- ✓ منطقة قبول H_0 هي $\left[-Z_{1-\frac{\alpha}{2}} , +Z_{1-\frac{\alpha}{2}}\right]$ بالنسبة للاختبار من طرفين ومنطقة الرفض هي خارج هذا المجال ($Z_{1-\frac{\alpha}{2}}$ تمثل القيمة المجدولة من التوزيع الاحتمالي الطبيعي إذا كان حجم العينة كبيرا (n>30)، وتمثل القيمة المجدولة من التوزيع الاحتمالي لستيودنت $t_{n-1}^{\frac{\alpha}{2}}$ إذا كان حجم العينة صغيرا (n≤30)).
- ✓ منطقة قبول H_0 هي $[Z_{1-\alpha} , +\infty]$ بالنسبة للاختبار من طرف واحد على اليمين ومنطقة الرفض هي خارج هذا المجال ($Z_{1-\alpha}$ تمثل القيمة المجدولة من التوزيع الاحتمالي الطبيعي إذا كان حجم العينة كبيرا (n>30)، وتمثل القيمة المجدولة من التوزيع الاحتمالي لستيودنت $t_{n-1}^{\acute{a}}$ إذا كان حجم العينة صغيرا (n≤30)).

- ✓ منطقة قبول H_0 هي $[-\infty, Z_{1-\alpha}]$ بالنسبة للاختبار من طرف واحد على اليسار ومنطقة الرفض هي خارج هذا المجال (القيم المجدولة هي نفسها القيم المجدولة المستخدمة بالنسبة للاختبار من طرف واحد على اليمين).

ملاحظة: إذا كانت قيمة δ^2 غير معلومة نقوم باستخدام مقدرتها غير المتحيزة $\hat{s}^2$ حيث:

$$\hat{s}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

أو $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ إذا كانت (n>30).

2.3.13. الدالة الاختبارية للنسبة P

مع العلم أن:

$P = \frac{R}{N}$، حيث R تمثل عدد الأفراد المتصفين بصفة معينة من بين مجتمع حجمه N.

$\hat{P} = \frac{r}{n}$، حيث r تمثل عدد الأفراد المتصفين بصفة معينة من بين عينة حجمها n.

و$\hat{\delta}^2_{\hat{P}} = \frac{\hat{P}\hat{Q}}{n}$ ، حيث: $\hat{Q}$ =1-$\hat{P}$.

وللقيام باختبار الفرضيات عند مستوى ثقة (1-α)، نقوم باحتساب الدالة الاختبارية $Y = \frac{\hat{P}-P}{\hat{\delta}_{\hat{P}}}$ ونرى إن كانت تنتمي لمنطقة القبول (zone d'acceptation) أو الرفض:[8]

- ✓ منطقة قبول H_0 هي $\left[-Y_{1-\frac{\alpha}{2}}, +Y_{1-\frac{\alpha}{2}}\right]$ بالنسبة للاختبار من طرفين ومنطقة الرفض هي خارج هذا المجال ($Y_{1-\frac{\alpha}{2}}$ تمثل القيمة المجدولة من التوزيع الاحتمالي الطبيعي إذا كان حجم العينة كبيرا (n>30)، وتمثل القيمة المجدولة من التوزيع الاحتمالي لستيودنت $t^{\frac{\alpha}{2}}_{n-1}$ إذا كان حجم العينة صغيرا (n≤30)).
- ✓ منطقة قبول H_0 هي $[Y_{1-\alpha}, +\infty]$ بالنسبة للاختبار من طرف واحد على اليمين ومنطقة الرفض هي خارج هذا المجال ($Y_{1-\alpha}$ تمثل القيمة المجدولة من التوزيع الاحتمالي الطبيعي إذا كان حجم العينة كبيرا (n>30)، وتمثل القيمة المجدولة من التوزيع الاحتمالي لستيودنت t^{α}_{n-1} إذا كان حجم العينة صغيرا (n≤30)).
- ✓ منطقة قبول H_0 هي $[-\infty, Y_{1-\alpha}]$ بالنسبة للاختبار من طرف واحد على اليسار ومنطقة الرفض هي خارج هذا المجال (القيم المجدولة هي نفسها القيم المجدولة المستخدمة بالنسبة للاختبار من طرف واحد على اليمين).

[8] G.BAILLARGEON, J.RAINVILLE, "Statistique appliquée : tests statistiques, régression et corrélation", Édition Trois-Rivières, Québec 1979, Page 71.

سلسلة تمارين مقترحة حول المحور الأول:

تمرين 01:

يظهر التوزيع الإحتمالي P(X) للمتغير العشوائي X، والذي يمثل عدد مرتادي أحد المطاعم للغداء، كما يلي:

$$P(X=x)=\begin{cases}0,2 & si\ x=100\\ 0,3 & si\ x=110\\ 0,2 & si\ x=118\\ 0,2 & si\ x=120\\ 0,1 & si\ x=125\end{cases}$$

المطلوب:

- ماهو العدد المتوقع للمترددين على المطعم وقت الغداء ؟
- ما هي قيمة التباين والإنحراف المعياري للتوزيع ؟

تمرين 02:

ليكن لدينا فندق يحوي على 50 غرفة، معدل إمتلاءه خلال فصل الربيع هو 75%.

المطلوب:

- إذا كان X متغير عشوائي يمثل عدد الغرف المحجوزة خلال يوم معين، ما نوع المتغير العشوائي وماهي مختلف القيم التي يمكن أن يأخذها ؟
- ماهو إحتمال أن يكون الفندق محجوزا بالكامل ؟

تمرين 03:

- ماهو إحتمال الحصول على 4 صور في 6 رميات لقطعة نقد متزنة ؟
- ماهو إحتمال الحصول على 3 مرات الرقم 6 في 4 رميات لنرد متزن ؟

تمرين 04:

إذا كانت 1% من المصابيح الكهربائية المنتجة في أحد المصانع غير صالحة.

المطلوب:

- ماهو إحتمال الحصول على مصباح غير صالح من عينة من 10 مصابيح ؟
- ماهو إحتمال الحصول على مصباح غير صالح من عينة من 200 مصباح ؟ (بإستخدام توزيع ذي الحدين وكذا بواسون كتقريب لذي الحدين)

تمرين 05:

تشير الخبرة الماضية إلى حدوث حادثي مرور عند تقاطع ما أسبوعيا.

المطلوب:

- ماهو إحتمال حصول 4 حوادث خلال أسبوع ؟
- ماهو إحتمال عدم حصول أي حادث خلال أسبوع ؟
- ما القيمة المتوقعة والإنحراف المعياري للتوزيع ؟

تمرين 06:

لوحظ في إحدى المصالح الإدارية أن متوسط عدد مرات تعطل جهاز كومبيوتر خلال أسبوع هو 0,3.

المطلوب:

- ماهو إحتمال أن لا يتعطل الجهاز خلال الأسبوع ؟
- ماهو إحتمال أن يتعطل الجهاز مرتين خلال الأسبوع ؟
- ماهو إحتمال أن لا يتعطل الجهاز خلال أسبوعين ؟

تمرين 07:

أحسب إحتمال إختيار 3 وحدات سليمة في عينة من 15 وحدة أخذت عشوائيا وبدون إرجاع من مجموعة من 25 وحدة منها 10 وحدات سليمة.

تمرين 08:

- أحسب إحتمال إختيار سيدتين في عينة من 5 أفراد مسحوبة عشوائيا بدون إحلال من مجموعة من 9 أفراد منهم 4 سيدات.
- أحسب إحتمال إختيار 8 رجال في عينة من 10 أفراد مسحوبة عشوائيا بدون إحلال من مجتمع من 1.000 فرد نصفهم رجال (بإستخدام التوزيع فوق الهندسي وكذا ثنائي الحد كتقريب للتوزيع فوق الهندسي).

تمرين 09:

على إعتبار أن صلاحية الشاحنات في مؤسسة ما هي متغير عشوائي بأمل رياضي قدره 14 سنة.

- ما نوع المتغير العشوائي X وماهي مختلف القيم التي يمكن أن يأخذها ؟

- عين الوسيط λ، ثم أكتب دالة الكثافة لهذا المتغير ؟
- أحسب التباين ؟
- أحسب إحتمال أن تكون مدة الصلاحية لشاحنة من هذه المؤسسة أكبر من 20 سنة؟

تمرين 10:

إذا كان زمن الإنتظار في أحد شبابيك إدارة ما متغير عشوائي x (محسوب بالدقائق) بوسيط λ=0.08.

- ماهو إحتمال أن ينتظر شخص ما أقل من 10 د ؟
- ماهو إحتمال أن ينتظر شخص ما لمدة تتراوح بين 10 د و30 د على الأكثر ؟

تمرين 11:

أعط الصيغ التالية:

- إحتمال أن يقع المتغير المتصل X بين X_1، X_2.
- دالة الكثافة للتوزيع الطبيعي.
- القيمة المتوقعة والتباين للتوزيع الطبيعي.
- التوزيع الطبيعي المعياري أو القياسي.
- ماهو متوسط وتباين التوزيع الطبيعي القياسي.

تمرين 12:

أوجد المساحة تحت منحنى التوزيع الطبيعي القياسي بين:

- ✓ $z = \pm 2{,}58,\ z = \pm 1{,}96,\ z = \pm 1{,}64$.
- ✓ $z = 0{,}90$ و $z = 2{,}10$.
- ✓ إلى اليسار من $z = 0{,}90$.
- ✓ إلى اليمين من $z = 2{,}10$.
- ✓ إلى اليسار من $z = 0{,}90$ وإلى اليمين من $z = 2{,}10$.

تمرين 13:

متغير عشوائي X يتبع التوزيع الطبيعي مع $\mu = 67$ و$\delta = 3$، ما إحتمال أن هذا المتغير العشوائي سيأخذ القيمة بين:

67 و70/ 60 و70/ 60 و65/ أقل من 60/ أكبر من 65.

تمرين 14:

تتبع درجات إختبارات الذكاء لمتطوعي الجيش في سنة ما التوزيع الطبيعي مع $\mu = 110$ و$\delta = 10$. ويريد الجيش أن يعطي تدريبا متقدما لأعلى 25% في درجات إختبارات الذكاء.

ماهي أقل درجة في إختبارات الذكاء التي تقبل لحضور التدريب المتقدم ؟

المحور الثاني: جمع البيانات وطرق المعاينة

جمع البيانات الأولية يعتمد عادة على عملية الاستقراء أو الإستطلاع. هذه الأخيرة تشمل عينة من مجتمع الدراسة وتعتمد في نجاحها على النقاط التالية:

- ✓ تحديد الهدف من البحث بشكل واضح مع وضع الفروض الاحصائية المناسبة، وتعد هذه المرحلة أساسية قبل البدء بأي دراسة، إذ أن عدم وضوح أهداف البحث قد يقود إلى نتائج لا تتوافق مع الأهداف المتوخاة من هذه الدراسة. كما يقصد بالفرض الاحصائي بأنه تفسير مبدئي للظاهرة موضوع الدراسة، ويحتاج الفرض إلى بيانات يتم جمعها وتحليلها وفي ذلك يقرر الباحث إما قبول الفرض أو رفضه وبدأ البحث عن الفرض البديل في ضوء البيانات المتاحة للباحث والتي تم جمعها عن الظاهرة موضوع الدراسة. ومن خلال الفروض يتم تحديد الهدف من الدراسة وتحديد الجداول الاحصائية اللازمة[9].
- ✓ إختبار الاستبيان المعد للمسح، وقد يكون من نتائج هذا الاختبار بعض التحسينات على الاستبيان بالاضافة إلى التعرف بشكل فعلي على التكلفة، والتي قد تكون مختلفة عما هو متوقع[10].
- ✓ تحديد مجتمع الدراسة من خلال الأطر الاحصائية، وتعد هذه المرحلة أساسية لما يجب من توفير إطار جيد لمفردات المجتمع قبل الشروع في أخذ العينة اللازمة للبحث. كما يجب على هذه الأطر الاحصائية أن تغطي كامل مفردات المجتمع من دون نقصان أو إزدواجية.
- ✓ إختيار طريقة المعاينة المناسبة والمتوافقة مع موضوع الدراسة.
- ✓ تحديد حجم العينة اللازم سحبها لتقدير جيد لخصائص مفرداتها وبالتالي تعميم النتائج المتوصل إليها على كامل المجتمع.

1. مصطلحات أساسية

- ✓ **المجتمع (Population):** ويشمل جميع المفردات الحاملة للمعلومات المطلوبة ذات العلاقة بالمشكلة المدروسة والتي يسعى الباحث إلى أن يعمم النتائج عليها.
- ✓ **العينة (Échantillon):** وتمثل مجموعة جزئية مختارة من مفردات المجتمع لتشملها عملية الاستقراء، ويعتمد هذا الاختيار على مدى درجة تمثيل العينة لخصائص المجتمع الأصل.
- ✓ **الوحدة الاحصائية (Unité statistique):** وتكون إما مفردة (شخص، شيء، حيوان،...) وإما مجموعة تحوي على عدة مفرادت (دولة، بلدية، عائلة،...)، ومجموع هذه الوحدات الاحصائية يشكل مكونات كل من المجتمع الأصلي والعينة المسحوبة.

[9] جابر أحمد بسيوني، مرجع سبق ذكره، ص17.
[10] إبراهيم علي، "مدخل في تقنيات المعاينة الإحصائية"، مجلس التخطيط القطري، الدوحة 2007، ص18.

- ✓ **الإطار الإحصائي (Cadre d'échantillonnage):** ويمثل مصدر الحصول على كافة مفردات المجتمع الأصلي مثل: قائمة داخلية لهيئة معينة، قائمة أرقام الهاتف، مخطط المدينة، ...

2. أسلوب المعاينة

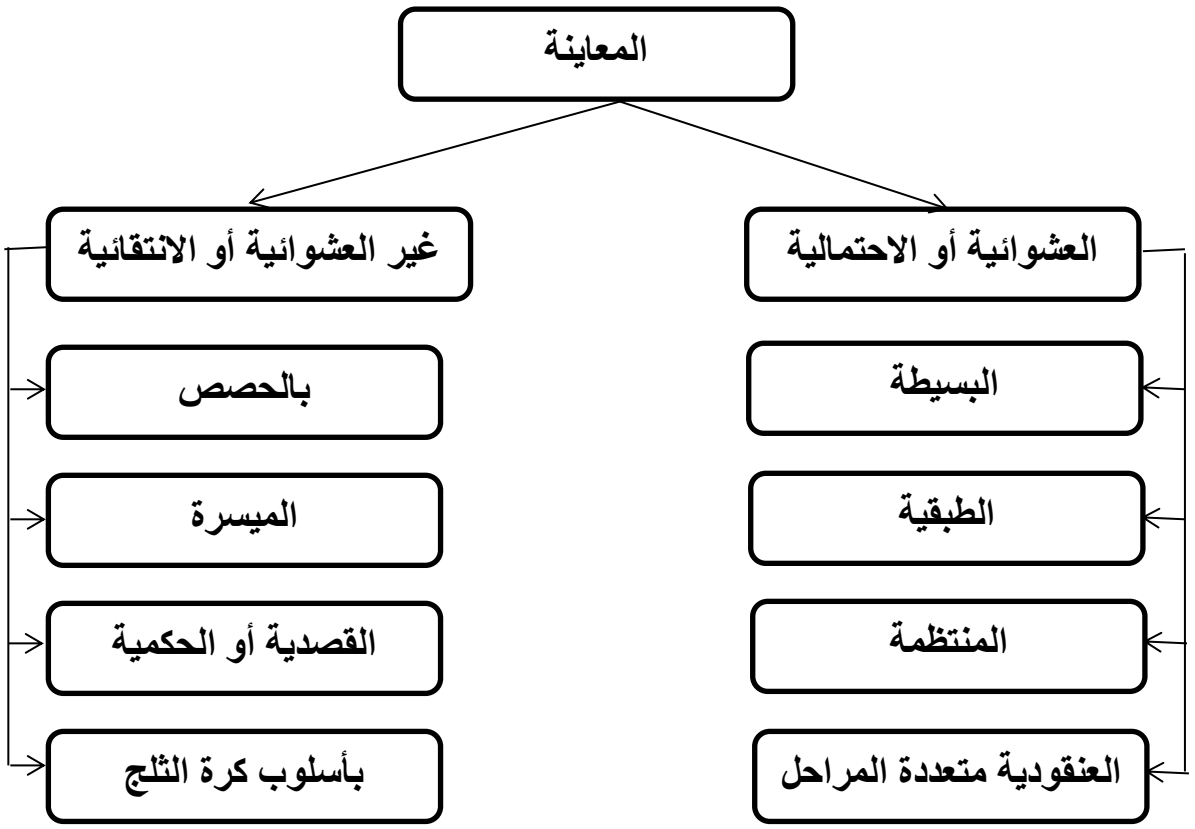
تعتبر العينة جزء من المجتمع يتم اختيارها بطريقة عشوائية وغير متحيزة لكي يتم دراستها للتعرف على خصائص المجتمع الأصل الذي سحبت منه. ويتميز أسلوب العينة بعدد من المميزات منها:

- ✓ يعطي نتائج سريعة بسبب سرعة الحصول على البيانات وتحليلها.
- ✓ يوفر الوقت، الجهد والتكاليف.
- ✓ بالنسبة لبعض الدراسات تعتبر العينات الأسلوب الاحصائي الوحيد القابل للتطبيق مثل عمل تجارب دواء جديد على بعض الحيوانات وغيره من الأمثلة العديدة.

3. طرق المعاينة

تختلف أنواع العينات تبعا لاختلاف خصائص المجتمع المراد دراسته، وبصفة عامة يمكن التمييز بين قسمين من العينات يمكن تلخيصهما من خلال الشكل الآتي:

شكل رقم 2-1: أساليب المعاينة



المصدر: من إعداد الأستاذ

وفيما يلي شرح لمختلف هذه العينات:

أولا: العينات الاحتمالية (Échantillons probabilistes ou aléatoires)

ويمكن تعريفها بأنها العينات التي يتم اختيار مفرداتها بأسلوب يوفر لكل وحدة من وحدات المعاينة بمجتمع الدراسة احتمالا ثابتا ومحددا لاختيار العينة بحيث لا يتدخل الباحث في عملية الاختيار بل يتم عشوائيا، ويمكن من خلال العينات الاحتمالية حساب أخطاء المعاينة والاستدلال الاحصائي[11] وكذا تعميم نتائجها على المجتمع ككل.

بالرغم من كل هذه المزايا فإن هذه الأخيرة تتطلب توفر قائمة لكل مفرادت المجتمع موضوع الدراسة من خلال أطر إحصائية ملائمة قبل الشروع في اختيار العينة، وهذا ما لا يمكن أن يتوفر في كل حالات.

وفيما يلي الأنواع الأربع من العينة الاحتمالية التي يمكن القيام بها:

أ. العينة العشوائية البسيطة (Échantillon aléatoire simple)

وعلى اعتبار هذا النوع من العينات من أبسط الأنواع فإن طريقة المعاينة العشوائية البسيطة تعد من الطرق الأساسية في تقنيات المعاينة والتي تبنى عليها الطرق الأخرى.

ويقصد بالعشوائية بأنه يكون لكل فرد من مفردات المجتمع نفس فرصة الاختيار ضمن العينة أي نفس إحتمال الظهور (لأن المجتمع تام التجانس أي أن كل مشاهدة فيه لها نفس الخصائص والصفات بالنظر لموضوع البحث).

وتتم عملية الاختيار بأن توضع لكل مفردة رقم معين وتخلط الأرقام مع بعضها خلطا جيدا ثم يتم سحب عدد مفردات العينة بطريقة عشوائية بطريقة السلة أو بطريقة جداول الأرقام العشوائية أو باستخدام الحاسب الآلي.

وتعتبر العينة العشوائية البسيطة سهلة الاختيار في حالة المجتمعات الصغيرة ويصعب استخدامها في المجتمعات الطبقية حيث لا تضمن تمثيل كل طبقة من هذه الطبقات بنفس نسبتها في المجتمع.

ونميز عادة بين نوعين من المعاينة العشوائية البسيطة (السحب)، مع الإعادة وبدون إعادة. فعدد العينات بحجم (n) الممكن سحبها من مجتمع حجمه (N) من دون إعادة هو:

$$C_N^n = \frac{N!}{n!(N-n)!}$$ (التوفيقة بدون إعادة)

[11] جابر أحمد بسيوني، مرجع سبق ذكره، ص22.

في المقابل، فإن عدد العينات بحجم (n) الممكن سحبها من مجتمع حجمه (N) بالإعادة هو:

$$\alpha_N^n = N^n$$ (الترتيبة مع الإعادة)

مثال 2-1:

إذا أردنا إختيار فوجين من قسم سنة ثانية ماستر تخصص تسويق مصرفي بكلية الاقتصاد بجامعة الجلفة وذلك لإجراء دراسة معينة، فماهو عدد العينات التي يمكن اختيارها من بين 04 أفواج؟

الحل:

لتكن رموز الأفواج: 1، 2، 3، 4.

الاختيار لا يكون هنا إلا بدون إعادة، وعليه يكون عدد العينات الممكن اختيار:

$$C_4^2 = \frac{4!}{2!.(4-2)!} = 6$$

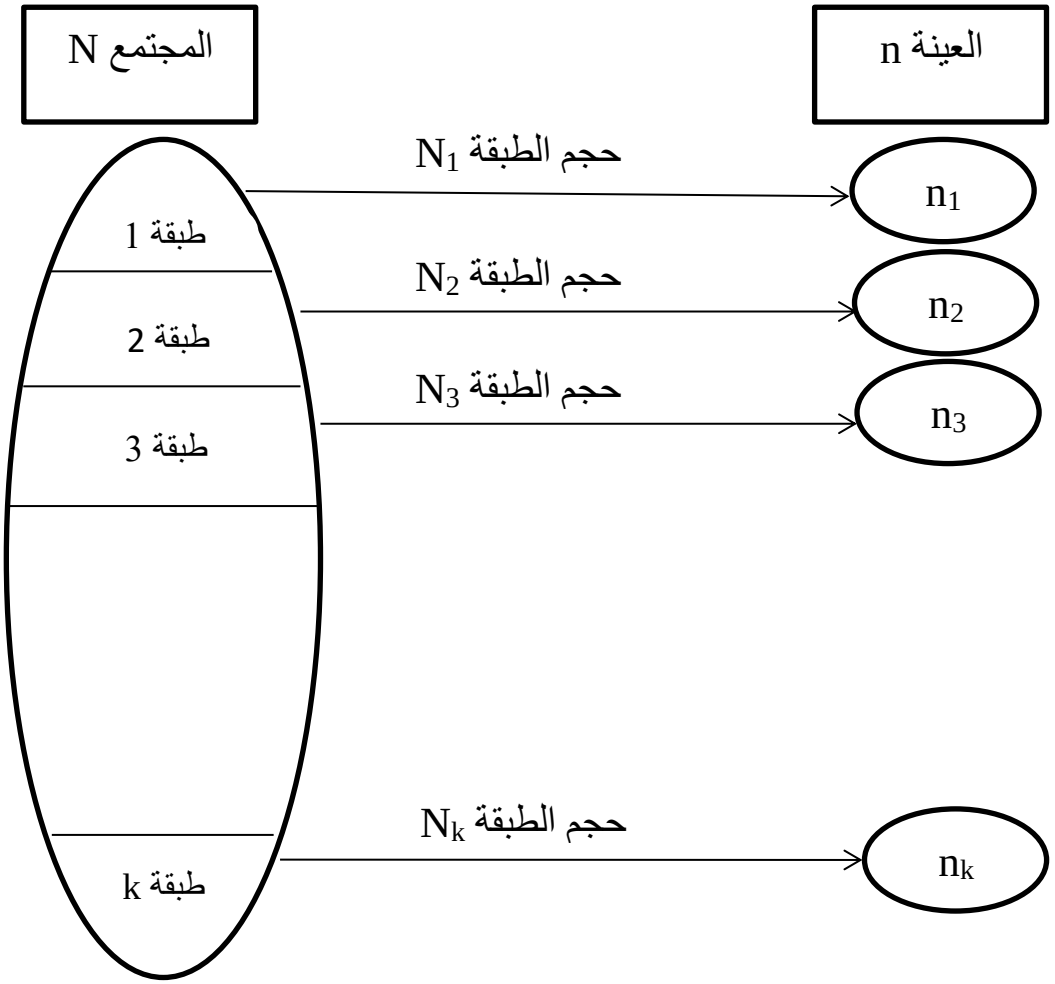
وهي الموضحة فيما يلي:

(1,2), (1,3), (1,4), (2,3), (2,4), (3,4)

ب. العينة العشوائية الطبقية (échantillon aléatoire stratifié)

المجتمع في هذه الحالة مقسم إلى مجموعات وذلك وفق معيار معين (الفئة العمرية، المنصب الإداري في الشركة، ...)، كل مجموعة تسمى طبقة (strate) وكل طبقة تامة التجانس في مفرداتها. نقوم بسحب عينة عشوائية بسيطة من كل طبقة مع الأخذ بعين الاعتبار نسبة كل طبقة بالمقارنة مع الحجم الكلي للمجتمع، ليكون مجموع العينات المسحوبة من كل طبقة هو العينة الطبقية العشوائية المطلوبة. بالنسبة لهذا الأسلوب من المعاينة، مفردات المجتمع لا تملك نفس احتمال الظهور مما يستوجب تجانس الطبقات (أنظر الشكل رقم 06).

شكل رقم 2-2: العينة العشوائية الطبقية



المصدر: من إعداد الأستاذ

حيث أن:

$$N = N1+N2+ \dots +Nk = \sum_{i=1}^{k} N_i$$

$$n = n1+n2+ \dots +nk = \sum_{i=1}^{k} n_i$$

وأن حجم العينة العشوائية البسيطة المسحوبة من الطبقة (i) يحدد بالعلاقة التالية:

$$n_i = n \cdot \frac{N_i}{N}$$

مثال 2-2:

على إفتراض أن 55% من طلبة جامعة الجلفة هم إناث و45% ذكور، وأردنا سحب عينة عشوائية طبقية بحجم 140 طالب، المطلوب:

- ✓ وضح كيفية سحب هذه العينة.

الحل:

- حجم العينة العشوائية البسيطة المسحوبة من الطلبة الإناث: 140x55%=77 طالبة.
- حجم العينة العشوائية البسيطة المسحوبة من الطلبة الذكور: 140x45%=63 طالب.

نلاحظ أن هذه الطريقة تسمح بتحقيق التناسب الطبقي المطلوب في العينة وذلك بالمقارنة مع ما يميز المجتمع. في المقابل ولاستعمال هذه الطريقة يجب توفر معلومات عن التوزيع الطبقي للمجتمع.

ج. العينة العشوائية المنتظمة (échantillon aléatoire systématique)

إذا كان مجتمع الدراسة يتصف بأنه يمكن ترتيبه بشكل تصاعدي أو تنازلي تكون العينة المناسبة للسحب هي العينة المنتظمة، ومثال على ذلك إذا أردنا سحب عينة من دفتر الشيكات فإن الشيكات مرتبة ترتيبا تصاعديا يمكن الاعتماد عليه في سحب عينة منتظمة[12].

لذا فإن هذه العينة تتطلب ترقيم مشاهدات المجتمع ثم اختيار مشاهدة عشوائيا من أول $(k = \frac{N}{n})$ من المشاهدات، وإذا أردنا اختيار المشاهدة الثانية فإننا ننتقل إلى:

$$k + \frac{N}{n}$$

أي أن مقدار الفترة بين المشاهدة الأولى والمشاهدة الثانية هو $\frac{N}{n}$ وهكذا نضيف $\frac{N}{n}$ للمشاهدة الثانية لنجد المشاهدة الثالثة وهكذا بالنسبة لبقية مشاهدات العينة.

مثال 2-3:

المطلوب اختيار عينة عشوائية منتظمة حجمها n=10 من مجتمع حجمه N=200 رقمت أفراده من 1 إلى 200.

الحل:

يمكن تقسيم المجتمع إلى $k = \frac{N}{n} = \frac{200}{10} = 20$ مجموعة، والذي يمثل أيضا طول الفترة.

ثم نقوم باختيار مفردة واحدة من المجموعة الأولى بطريقة عشوائية بسيطة وبفرض اختيار الرقم (4) ثم نقوم بتحديد باقي مفردات العينة بإضافة طول الفئة (20) على الرقم (4) فنحصل على 10 مفردات بالأرقام 4، 24، 44، 64، 84، 104، 124، 144، 164، 184 والتي تمثل مفردات العينة المختارة.

ويلاحظ أن هذا الأسلوب في اختيار العينات يحتاج إلى تكاليف وجهد أقل من العينة العشوائية البسيطة.

د. العينة العنقودية متعددة المراحل (échantillon par grappes à plusieurs degrés)

يتم اختيار العينة وفق هذا الأسلوب على عدة مراحل، ويستخدم هذا الأسلوب في حالة ما إذا كان حجم المجتمع كبيرا جدا فيتم تقسيمه إلى وحدات تسمى بالوحدات الأولية، ثم يتم اختيار عينة عشوائية بسيطة من هذه الوحدات الأولية، ونقوم كذلك بتقسيم الوحدات المختارة إلى وحدات ثانوية ثم اختيار عينة عشوائية بسيطة منها وهكذا نستمر بالتقسيم والاختيار إلى أن نصل إلى العينة الأخيرة التي تحوي على مفردات تستنبط منها المعلومات المطلوبة.

[12] ثائر فيصل شاهر، مرجع سبق ذكره، ص20.

مثال 2-4:

من أجل دراسة واقع قطاع الفلاحة بالمناطق السهبية في الجزائر، يتم اختيار عدد من الولايات الواقعة بهذه المناطق، ثم عدد من الدوائر بالنسبة لكل ولاية مختارة ثم يتم اختيار عدد من البلديات داخل كل دائرة، لتكون الدراسة منصبة على واقع النشاط بهذه البلديات المختارة كمفردات للعينة متعددة المراحل.

ثانيا: العينات غير الاحتمالية (Échantillons non probabilistes ou raisonnés)

المعاينة غير الاحتمالية تتعمد اختيار مفردات معينة أي أنه تتميز بالانتقائية، ولذلك فهي تعرف أحيانا بأنها العينة العمدية أو التحكمية.

هنالك دراسات يصعب تحديد المجتمع الأصلي لها مثل دراسة أحوال المنحرفين أو المدمنين أو المتهربين من دفع الضرائب...، فهذه المجتمعات محددة ولكن أفرادها ليسوا معروفين وهذا ما يفسر استحالة اللجوء إلى المعاينة العشوائية، فيعتمد الباحث إلى أسلوب العينة غير العشوائية.

وبالنسبة للمعاينة غير الاحتمالية لا يمكن احتساب أخطاء المعاينة أو الاستدلال الاحصائي أو حتى تعميم النتائج المتوصل إليها على كافة المجتمع. فيلجأ الباحث لأن يذكر أن نتائج تحليله للعينة القصدية، تخص وحدات العينة المختارة فقط وليس مجتمع الدراسة، كأن يقول "يرى أفراد العينة، يؤثر البرنامج محل الدراسة على أفراد العينة، ...". ومن أهم أنواع هذه العينات نجد:

أ. **المعاينة بالحصص (échantillonnage par quotas)**

يستخدم هذا الأسلوب غالبا في دراسات السوق وكذا الدراسات المهتمة بموقف الرأي العام من قضايا معينة، وفيها يتم تحديد حصص معينة تمثل مجتمع الدراسة تمثيلا نسبيا، ويعتمد تحديد الحصص على توفر معلومات دقيقة حول توزيع المجتمع بالنظر لعامل معين من العوامل والتي تسمى بمتغيرات المراقبة (الجنس، الفئة العمرية، التخصص الجامعي، ...)، ويتم اختيار مفردات العينة من داخل كل قسم بطريقة عمدية غير عشوائية، وهذا ما يمثل وجه الاختلاف بين هذه الأسلوب من المعاينة والمعاينة العشوائية الطبقية.

مثال 2-5: المعاينة بالحصص

	المجتمع	العينة
ذكور	%51,17	511
إناث	%48,83	489
المجموع	%100	1000

مثال 2-6: المعاينة بالحصص المتقاطعة

المجتمع	20-34سنة	35-49سنة	50-60سنة	المجموع
ذكور	%15,25	%20,37	%12,77	%48,39
إناث	%20,84	%17,54	%13,23	%51,61
المجموع	%36,09	%37,91	%26,00	%100

المجتمع	20-34سنة	35-49سنة	50-60سنة	المجموع
ذكور	152	204	128	484
إناث	209	175	132	516
المجموع	361	379	260	1000

ب. العينة الميسرة (échantillonnage de convenance ou de commodité)

ويكون فيها إختيار مفردات العينة وفق معايير عملية متمثلة أساسا في مدى إمكانية الوصول إليهم، مدى تفرغهم للعملية وكذا إقناعهم على المشاركة.

ج. العينة الحكمية أو القصدية (échantillonnage selon le jugement)

يختار الباحث هذه العينة بشكل حر وشخصي يرى أنه سيمكنه من تحقيق أهداف البحث (يعتمد اختيار الأفراد على مهارات الفرد أو تمثيله)، فإذا أراد مثلا باحث أن يدرس أثر تطبيق المعايير المحاسبية IAS/IFRS على قطاع أو مؤسسة اقتصادية معينة فإنه يختار عددا من المحاسبين كعينة قصدية تحقق أغراض دراسته.

د. المعاينة بأسلوب كرة الثلج (échantillonnage boule de neige)

يستخدم هذا الأسلوب من المعاينة في حالة ندرة وحدات المعاينة أو عدم وجود إطار معاينة سليم من النقائص. نطلب من الشخص المجيب أن يشير إلى شخص آخر له نفس خصائصه يمكن أن نستجوبه، وهكذا... (يتم بناء العينة تدريجيا بناء على إقتراح من الأفراد الذين شملهم الاستطلاع).

سلسلة تمارين مقترحة حول المحور الثاني:

تمرين 01:

إذا أردنا اختيار عينة مكونة من شخصين من مجتمع يحوي خمسة أشخاص، وكان لكل شخص نفس احتمال الظهور ضمن مفردات العينة. المطلوب:

1/ ما نوع المعاينة في هذه الحالة؟

2/ ما نوع السحب في هذه الحالة؟

3/ كم عينة يمكن تشكيلها في هذه الحالة؟ وماهي هذه العينات إذا ما قمنا بترميز الأشخاص الخمس بالحروف من A إلى E؟

تمرين 02:

إذا كان لدينا في أحد المصانع (25) إداري و(75) عامل مكلف بالإنتاج، وأردنا سحب عينة بحجم (20) عامل. المطلوب:

1/ ما نوع المعاينة في هذه الحالة؟

2/ ما نوع السحب في هذه الحالة؟

3/ كيف تتم عملية السحب؟

تمرين 03:

إذا أردنا سحب عينة بحجم 05 صكوك من دفتر شيكات أحد الزبائن لدى بنك ما لأجل دراسة مستوى إنفاقه الفردي وكانت الشيكات مرقمة من 1 إلى 50. المطلوب:

1/ ما نوع المعاينة في هذه الحالة؟

2/ ما نوع السحب في هذه الحالة؟

3/ كيف تتم عملية السحب؟

تمرين 04:

حدد نوع العينة الاحتمالية المستعملة في كل حالة من الحالات التالية:

1/ اختيار طلبة من جامعة الجلفة للمشاركة في وضع برنامج لمختلف التظاهرات الرياضية للموسم الجامعي الحالي.

2/ دراسة واقع الانفاق الصحي للعائلات في عدة ولايات جزائرية.

3/ دراسة الانفاق العائلي والمطلوب سحب عينة حجمها (n=20) من مجتمع حجمه (N=200) عائلة مرقمة من (1-200).

4/ سحب عينة لدراسة إمكانية تصدير الفائض من محصول التمور.

تمرين 05:

حدد نوع العينة غير الاحتمالية المستعملة في كل حالة من الحالات التالية:

1/ لدراسة مدى رضى الزبائن بالخدمات المقدمة من قبل مركز تجاري معين، تم استجواب مئة أول زبون دخلوا المركز التجاري.

2/ لدراسة تطلعات الرياضيين إزاء منتوج رياضي معين، قام المصنع باستجواب زملاء له رياضيين كما طلب منهم مساعدته للتواصل مع رياضيين آخرين.

3/ لدراسة دور التأمين على القروض في دعم الائتمان المصرفي، تم سحب عينة مكونة من مجموعة من موظفي القطاع المصرفي.

4/ سحب عينة مكونة من 40 إمراة من أجل الموضوع محل الدراسة: منهن 20 من الوسط الريفي (15 سنهن 25 سنة و05 أكبر من ذلك)، و20 إمراة من الوسط الحضري (10 منهن سنهن 25 سنة و10 أكبر من ذلك).

المحور الثالث: العينة العشوائية البسيطة

سنتناول من خلال هذا المحور مختلف العمليات المتعلقة بمتوسط المجتمع μ، المجموع الكليT، النسبة P وكذا عدد المفردات المتصفة بصفة معينة R.

1. تقدير متوسط المجتمع μ والمجموع الكلي T

إذا كان المتغير العشوائي x يتبع التوزيع الطبيعي بمتوسط μ وتباين δ^2 (μ وδ^2 معلمتين مجهولتين). ولتقديرهما نقوم بسحب العينة العشوائية البسيطة x1, x2, …, xn حيث n تمثل حجم العينة.

الوسط الحسابي للعينة $\bar{x}$ هو تقدير غير متحيز لمتوسط المجتمع μ $(E(\bar{x}) = \mu)$ ويتصف بكافة خصائص المقدر الاحصائي الجيد لذلك سيستخدم كتقدير لمتوسط المجتمع μ حيث:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

وتباين معطى بالعلاقة: $V(\bar{x}) = \frac{\delta^2}{n}$.

وبما أنs^2 يعتبر مقدرة غير متحيزة لتباين المجتمع δ^2 حيث:[13]

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

أو

$$s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

إذا كانت n كبيرة (n>30)، يصبح تباين متوسط المجتمع المقدر $\bar{x}$ كما يلي: $\widehat{V(\bar{x})} = \frac{s^2}{n}$ إذا كانت المعاينة من مجتمع غير محدود وبسحب مع الإعادة.

في المقابل، عندما تكون المعاينة من مجتمع محدود حجمه N وكانت قيمة δ^2 معلومة وبسحب من دون إعادة فإن:

$$=(\frac{N-n}{N-1})\frac{\delta^2}{n}\delta_{\bar{x}}^2$$

أما عندما تكون المعاينة من مجتمع محدود حجمه N وكانت قيمة δ^2 مجهولة وبسحب من دون إعادة فإن (أنظر الجدول رقم 3-1):

[13] حسين علوان مطلق، "جمع البيانات وطرق المعاينة"، العبيكان للنشر، الظهران 2010، ص366.

الجدول رقم 3-1: القيم الفعلية والمقدرة للمتوسط μ والمجموع الكلي T

	المتوسط	المجموع الكلي
المعلمة الحقيقية (المجتمع)	$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$	$T=N\mu$
المقدرة	$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$	$\hat{T} = N\bar{x}$
التباين	$\delta_{\bar{x}}^2 = (1 - \frac{n}{N})\frac{\delta^2}{n}$	$\delta_{\hat{T}}^2 = N^2(1 - \frac{n}{N})\frac{\delta^2}{n}$
التباين المقدر	$\hat{\delta}_{\bar{x}}^2 = (1 - \frac{n}{N})\frac{s^2}{n}$	$\hat{\delta}_{\hat{T}}^2 = N^2(1 - \frac{n}{N})\frac{s^2}{n}$
الانحراف المعياري	$\hat{\delta}_{\bar{x}} = \sqrt{(1 - \frac{n}{N})\frac{s^2}{n}}$	$\hat{\delta}_{\hat{T}} = N\sqrt{(1 - \frac{n}{N})\frac{s^2}{n}}$

المصدر: من إعداد الأستاذ

2.مجال الثقة بالنسبة لمتوسط المجتمع μ والمجموع الكلي T

1.2. مجال الثقة لمتوسط المجتمع μ

إذا كان المتغير x يتبع توزيع احتمالي معين بمتوسط μ وتباين δ^2، ومن خلال نظرية النهاية المركزية (TCL) فإن:

$Z = \frac{\bar{x}-\mu}{\hat{\delta}_{\bar{x}}}$ يتبع التوزيع الاحتمالي الطبيعي بمتوسط معدوم وتباين يساوي الواحد $Z \sim N(0,1)$ عندما يكون حجم العينة كبيرا (n>30).

مع العلم أن:

$$P(-z \leq Z \leq +z) = 1 - \alpha$$

فإنه يمكن تشكيل مجال الثقة لمتوسط المجتمع μ غير المعلوم بمستوى معنوية α كما يلي:

لدينا: $-z \leq Z \leq +z$

$$-z \leq \frac{\bar{x} - \mu}{\hat{\delta}_{\bar{x}}} \leq +z$$

$$\bar{x} - z.\hat{\delta}_{\bar{x}} \leq \mu \leq \bar{x} + z.\hat{\delta}_{\bar{x}}$$

أو $\mu \in \left[\bar{x} - z.\hat{\delta}_{\bar{x}} , \bar{x} + z.\hat{\delta}_{\bar{x}}\right]$

وعليه يكون مجال الثقة بالنسبة ل μ بمستوى ثقة $(1-\alpha)$ هو:

$$\mu \in \left[\bar{x} - z_{1-\frac{\alpha}{2}}.\hat{\delta}_{\bar{x}} , \bar{x} + z_{1-\frac{\alpha}{2}}.\hat{\delta}_{\bar{x}}\right]$$

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا $(n\leq 30)$، نستعمل توزيع ستيودنت (Student) بدرجة حرية $(n-1)$ أي $Z \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

مثال 3-1:

لأجل دراسة الاستهلاك الشهري للكهرباء للأسرة الواحدة في إحدى البلديات المكونة من 1000 أسرة، قمنا بسحب عينة عشوائية بسيطة حجمها $n=100$ أسرة، وإذا علمت أن الوسط الحسابي للإستهلاك لهذه العينة هو $\bar{x}=22$ واط وتباين $s^2 = 2$.

المطلوب:

1/ أوجد مجال الثقة بالنسبة لمتوسط الاستهلاك الشهري الأسري في هذه البلدية μ وذلك عند مستوى ثقة 0.95.

2/ إختبر الفرضية القائلة بأن متوسط الاستهلاك الشهري الأسري في هذه البلدية μ لا يزيد عن 21 واط وذلك عند مستوى ثقة 0.95.

الحل:

1/ من خلال المعطيات يمكن حساب مقدرة تباين متوسط الاستهلاك الشهري الأسري $\hat{\delta}_{\bar{x}}$:

$$\hat{\delta}_{\bar{x}}^2 = \left(1 - \frac{n}{N}\right)\frac{s^2}{n} = \left(1 - \frac{100}{1000}\right)\frac{2}{100} = 0{,}018$$

الانحراف المعياري يساوي: $\hat{\delta}_{\bar{x}} = \sqrt{0.018} = 0{,}134$.

من جدول التوزيع الطبيعي $N(0,1)$ نجد أن: $z_{1-\frac{\alpha}{2}} = 1{,}96$.

وعليه فإن مجال ثقة μ وذلك عند مستوى ثقة 0.95 هو:

$$\mu \in [22 - (1{,}96).(0{,}134) , 22 + (1{,}96).(0{,}134)]$$

$$\mu \in [21.73 , 22.26]$$

أي أن قيمة μ لا يمكن أن تقل عن 21.73واط ولا أن تزيد عن 22.26واط وذلك عند مستوى ثقة 0.95.

2/ الفرض القائل بأن متوسط الاستهلاك الشهري الأسري في هذه البلدية μ لا يزيد عن 21 واط وذلك عند مستوى ثقة 0.95، يجعلنا أمام إختبار من طرف واحد على اليسار كالآتي:

$$\begin{cases} H_0 : \mu \leq 21 \\ H_1 : \mu > 21 \end{cases}$$

نقوم باحتساب الدالة الاختبارية $Z_{cal} = \frac{\bar{x}-\mu}{\hat{\delta}_{\bar{x}}} = \frac{22-21}{0,134} = 7,46$.

نلاحظ أن قيمة Z_{cal} لا تنتمي لمجال منطقة قبول H_0: $]-\infty , 1.65]$[14]، إذن نرفض H_0 وبالتالي فإن متوسط الاستهلاك الشهري الأسري في هذه البلدية μ يزيد عن 21 واط وذلك عند مستوى ثقة 0.95.

2.2. مجال الثقة للمجموع الكلي للمجتمع T

كما رأينا سابقا بالنسبة لمجال ثقة متوسط المجتمع μ، فإن مجال الثقة بالنسبة للمجموع الكلي للمجتمع T غير المعلوم بمستوى معنوية α يمكن حسابه كما يلي:

بما أن تقدير المجموع الكلي للمجتمع $\hat{T}$ هو متغير عشوائي يتبع التوزيع الطبيعي بمتوسط $N\mu$ (T) وتباين $\hat{\delta}^2_{\hat{T}}$ (μ وδ^2 غير معلومتين)، ومن خلال نظرية النهاية المركزية (TCL) فإن:

$Z' = \frac{\hat{T}-T}{\hat{\delta}_{\hat{T}}}$ يتبع التوزيع الاحتمالي الطبيعي بمتوسط معدوم وتباين يساوي الواحد $Z' \sim N(0,1)$ عندما يكون حجم العينة كبيرا (n>30).

مع العلم أن:

$$P(-z' \leq Z' \leq +z') = 1 - \alpha$$

أي:

$$-z' \leq Z' \leq +z'$$

$$-z' \leq \frac{\hat{T} - T}{\hat{\delta}_{\hat{T}}} \leq +z'$$

$$\hat{T} - z'.\hat{\delta}_{\hat{T}} \leq T \leq \hat{T} + z'.\hat{\delta}_{\hat{T}}$$

$$T \in \left[\hat{T} - z'.\hat{\delta}_{\hat{T}} , \hat{T} + z'.\hat{\delta}_{\hat{T}}\right]$$ أو

وعليه يكون مجال الثقة بالنسبة ل T بمستوى ثقة (1-α) هو:

$$\boxed{T \in \left[\hat{T} - z'_{1-\frac{\alpha}{2}}.\hat{\delta}_{\hat{T}} , \hat{T} + z'_{1-\frac{\alpha}{2}}.\hat{\delta}_{\hat{T}}\right]}$$

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا (n≤30)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z' \sim t^{1-\frac{\alpha}{2}}_{n-1}$، بدل التوزيع الطبيعي $N(0,1)$.

[14] من جدول التوزيع الطبيعي $N(0,1)$ في الملحق نجد أن: $z_{1-\alpha} = 1,65$.

مثال 3-2:
بالعودة لبيانات المثال 3-1، المطلوب:

1/ أوجد مجال الثقة بالنسبة لإجمالي الاستهلاك الشهري للأسر في هذه البلدية T وذلك عند مستوى ثقة 0.95.

2/ إختبر الفرضية القائلة بأن إجمالي الاستهلاك الشهري للأسر في هذه البلدية T لا يقل عن 23.000واط وذلك بمستوى ثقة 0.95.

الحل:

1/ من خلال المعطيات نحتسب:

$$\widehat{\mathrm{T}} = \mathrm{N}\bar{\mathrm{x}} = 1000.(22) = 22000,\ \hat{\delta}_{\widehat{\mathrm{T}}} = N.\hat{\delta}_{\bar{\mathrm{x}}} = 1000.(0{,}134) = 134.$$

وعليه فإن مجال ثقة T وذلك عند مستوى ثقة 0.95 هو:

$$\mathrm{T} \in [22000 - (1{,}96).(134)\ ,\ 22000 + (1{,}96).(134)]$$
$$\mathrm{T} \in [21737{,}36\ ,\ 22262{,}64]$$

أي أن قيمة T لا يمكن أن تقل عن 21737,36واط ولا أن تزيد عن 22262,64واط وذلك عند مستوى ثقة 0.95.

2/ الفرض القائل بأن إجمالي الاستهلاك الشهري للأسر في هذه البلدية T لا يقل عن 23000واط وذلك بمستوى ثقة 0.95، يجعلنا أمام إختبار من طرف واحد على اليمين كالآتي:

$$\begin{cases} \mathrm{H}_0 : \mu \geq 23000 \\ \mathrm{H}_1 : \mu < 23000 \end{cases}$$

نقوم باحتساب الدالة الاختبارية $\mathrm{Z_{cal}} = \frac{\widehat{\mathrm{T}} - \mathrm{T}}{\hat{\delta}_{\widehat{\mathrm{T}}}} = \frac{22000 - 23000}{134} = -7{,}46$.

نلاحظ أن قيمة $\mathrm{Z_{cal}}$ لا تنتمي لمجال منطقة قبول $\mathrm{H_0}$: $[1.65, +\infty[$[15]، إذن نرفض $\mathrm{H_0}$ وبالتالي فإن إجمالي الاستهلاك الشهري للأسر في هذه البلدية T يقل عن 23000 واط وذلك عند مستوى ثقة 0.95.

3. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ والمجموع الكلي للمجتمع T

إن لحجم العينة دورا مهما في تحديد دقة النتائج المتوصل إليها. فمن المعلوم أنه كلما زاد حجم العينة كلما كانت هذه النتائج أكثر دقة وقريبة أكثر من معالم المجتمع محل الدراسة. غير أن أخذ عينة كبيرة قد لا يكون ضمن الامكانات المتاحة للباحث، ولهذا فتحديد حجم العينة العشوائية البسيطة الممثلة جيدا للمجتمع يعتمد على العوامل التالية:

[15] من جدول التوزيع الطبيعي $\mathrm{N}(0,1)$ في الملحق نجد أن: $z_{1-\alpha} = 1{,}65$.

1/ تحديد مستوى الدقة المناسب (مقدار الخطأ المسموح به)، والذي يعتمد على أهمية القرارات التي ستعتمد على نتائج هذه الدراسة.
2/ مستوى الثقة المطلوب في التقديرات التي سنحصل عليها (1-α) حيث α هو مستوى المعنوية[16].

1.3. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ

إذا كان x متغير عشوائي يتبع التوزيع الطبيعي بمتوسط μ وتباين δ^2. وأردنا تقدير حجم العينة n الواجب سحبه من مجتمع حجمه N لتقدير متوسط المجتمع μ، أن نحدد ما يلي:
1/ مستوى الدقة المناسب، حيث لا يزيد الفرق المطلق بين متوسط العينة ومتوسط المجتمع عن $d = |\bar{x} - \mu|$.
2/ مستوى الثقة (1-α).
3/ القيمة العيارية للوسط الحسابي $\bar{x}$ لها توزيع طبيعي قياسي بمتوسط معدوم وتباين يساوي الواحد، حيث:

$$Z = \frac{\bar{x} - \mu}{\delta_{\bar{x}}} \sim N(0\,,1)$$

ومنها فإننا نجد:

$$Z_{1-\frac{\alpha}{2}} = \frac{|\bar{x} - \mu|}{\delta_{\bar{x}}}$$

$$Z^2_{1-\frac{\alpha}{2}} = \frac{d^2}{\delta^2_{\bar{x}}}$$

$$Z^2_{1-\frac{\alpha}{2}} = \frac{d^2}{(1 - \frac{n}{N})\,\frac{\delta^2}{n}}$$

$$d^2 = \frac{\delta^2}{n}.Z^2_{1-\frac{\alpha}{2}} - \frac{\delta^2}{N}.Z^2_{1-\frac{\alpha}{2}}$$

$$\boxed{n = \frac{N.\delta^2}{\frac{Nd^2}{Z^2_{1-\frac{\alpha}{2}}} + \delta^2}}$$

ملاحظة:

- ✓ إذا كانت قيمة التباين δ^2 غير معلومة، نستخدم بدلها مقدرتها غير المتحيزة S^2.

[16] فتحي أحمد عاروري، "المعاينة الاحصائية - طرقها واستخداماتها"، الأكاديميون للنشر والتوزيع، عمان 2013، ص117.

- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

يمكن أن نلجأ أيضا لطريقة حد الخطأ في تقدير المتوسط لتحديد حجم العينة اللازم لتقدير متوسط المجتمع μ كما يلي:
باعتبار مقدار الخطأ في تقدير متوسط المجتمع B يساوي ضعف الإنحراف المعياري للمقدرة غير المتحيزة ل μ $(\bar{x})$، أي:

$$B = 2.\delta_{\bar{x}}$$

(تجدر الإشارة إلى أن مدى أي متغير عشوائي يبلغ تقريبا 4δ).
بتربيع الطرفين نجد:

$$B^2 = 4.\delta_{\bar{x}}^2$$

بتعويض $\delta_{\bar{x}}^2$ بعبارته نجد:

$$B^2 = 4.(1 - \frac{n}{N})\frac{\delta^2}{n}$$

$$D = \frac{B^2}{4} = \frac{N\delta^2 - n\delta^2}{nN}$$
$$n(ND + \delta^2) = N\delta^2$$

ومنه:

$$\boxed{n = \frac{N\delta^2}{ND + \delta^2}}$$

2.3. تقدير حجم العينة اللازم لتقدير المجموع الكلي للمجتمع T

مما سبق وبنفس الطريقة المعتمدة في تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ لدينا:

$$\hat{T} \sim N(T, \delta_{\hat{T}})$$
$$\frac{\hat{T} - T}{\delta_{\hat{T}}} \sim N(0,1)$$

ومنها فإننا نجد:

$$Z' = \frac{|\hat{T} - T|}{\delta_{\hat{T}}}$$

بوضع $d' = \left|\widehat{T} - T\right|$ وتربيع الطرفين نجد:

$$Z'^{2}_{1-\frac{\alpha}{2}} = \frac{d'^{2}}{\delta^{2}_{\widehat{T}}}$$

بتعويض $\delta^{2}_{\widehat{T}}$ بعبارته نجد:

$$Z'^{2}_{1-\frac{\alpha}{2}} = \frac{d'^{2}}{N^{2}.(1-\frac{n}{N})\frac{\delta^{2}}{n}}$$

$$d'^{2} = N^{2}.\frac{\delta^{2}}{n}.Z'^{2}_{1-\frac{\alpha}{2}} - N.\delta^{2}Z'^{2}_{1-\frac{\alpha}{2}}$$

ومنه فإن حجم العينة اللازم سحبه لتقدير المجموع الكلي للمجتمع T بمستوى دقة d' ومستوى معنوية α يعطى بالعلاقة:

$$\boxed{n = \frac{N^{2}.\delta^{2}}{\frac{d'^{2}}{Z'^{2}_{1-\frac{\alpha}{2}}} + N.\delta^{2}}}$$

ملاحظة:

- ✓ إذا كانت قيمة التباين δ^{2} غير معلومة، نستخدم بدلها مقدرتها غير المتحيزة S^{2}.
- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z' \sim t^{1-\frac{\alpha}{2}}_{n-1}$، بدل التوزيع الطبيعي $N(0,1)$.

يمكن أيضا أن نلجأ لطريقة حد الخطأ في تقدير المتوسط لتحديد حجم العينة اللازم لتقدير المجموع الكلي للمجتمع T كما يلي:

باعتبار مقدار الخطأ في تقدير المجموع الكلي للمجتمع B' يساوي ضعف الإنحراف المعياري لمقدرة T $(\widehat{T})$، أي:

$$B' = 2.\delta_{\widehat{T}}$$

بتربيع الطرفين نجد:

$$B'^{2} = 4.\delta^{2}_{\widehat{T}}$$

بتعويض $\delta^{2}_{\widehat{T}}$ بعبارته نجد:

$$B'^{2} = 4.N^{2}.(1-\frac{n}{N})\frac{\delta^{2}}{n}$$

$$D' = \frac{B'^{2}}{4N^{2}} = \frac{N\delta^{2} - n\delta^{2}}{nN}$$

$$n(ND' + \delta^2) = N\delta^2$$

ومنه:

$$\boxed{n = \frac{N\delta^2}{ND' + \delta^2}}$$

مثال 3-3:

بالعودة لبيانات المثالين 3-1 و3-2، المطلوب:

1/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير متوسط الاستهلاك الشهري الأسري في هذه البلدية بحيث لا يتعدى الفرق المطلق بين قيمة μ الفعلية والمقدرة ال 1 واط، $d = |\bar{x} - \mu| = 1$، وذلك بمستوى ثقة 0.95 ؟

2/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها بحيث يكون الحد على الخطأ في تقدير متوسط الاستهلاك الشهري الأسري في هذه البلدية B=1 ؟

3/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير إجمالي الاستهلاك الشهري للأسر في هذه البلدية بحيث لا يتعدى الفرق المطلق بين قيمة T الفعلية والمقدرة ال 1.000 واط، $d' = |\hat{T} - T| = 1.000$، وذلك بمستوى ثقة 0.95 ؟

الحل:

1/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير μ بحيث لا يتعدى الفرق المطلق، $d = |\bar{x} - \mu| = 1$، وذلك بمستوى ثقة 0.95:

$$n = \frac{1000.(2)}{\frac{1000.(1)^2}{(1,96)^2} + 2} = 7,62$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 8$.

2/ حجم العينة العشوائية البسيطة اللازم سحبها بحيث لا يتعدى حد الخطأ في تقدير متوسط الاستهلاك الشهري الأسري في هذه البلدية B=1:

$$D = \frac{B^2}{4} = \frac{(1)^2}{4} = 0,25$$

$$n = \frac{1000.(2)}{1000.(0,25) + 2} = 7,93$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 8$.

3/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير T بحيث لا يتعدى الفرق المطلق، $d' = |\hat{T} - T| = 1.000$، وذلك بمستوى ثقة 0.95:

$$n = \frac{(1000)^2.(2)}{\frac{(1000)^2}{(1,96)^2} + 1000.(2)} = 7,62$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 8$.

4. تقدير حجم العينة اللازم لتقدير μ وT باستخدام مستوى الدقة النسبي

يمكن أن نقوم أيضا بتقدير حجم العينة اللازم سحبها من أجل تحقيق مستوى معين من الدقة النسبية لتقدير متوسط المجتمع μ أو لتقدير المجموع الكلي للمجتمع T.

1.4. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ بمستوى دقة نسبي p

يعطى مستوى الدقة النسبية لتقدير متوسط المجتمع μ بالعبارة:

$$p = \left|\frac{\bar{x} - \mu}{\mu}\right|$$

لدينا:

$$Z = \frac{\bar{x} - \mu}{\delta_{\bar{x}}}$$

ومنها فإننا نجد:

$$Z_{1-\frac{\alpha}{2}} = \frac{\left|\frac{\bar{x} - \mu}{\mu}\right|}{\frac{\delta_{\bar{x}}}{\mu}}$$

بتربيع الطرفين يصبح لدينا:

$$Z^2_{1-\frac{\alpha}{2}} = \frac{p^2}{(1 - \frac{n}{N})\frac{\delta^2}{n\mu^2}}$$

$$Z^2_{1-\frac{\alpha}{2}} = \frac{p^2}{\frac{N\delta^2 - n\delta^2}{nN\mu^2}}$$

$$nN\mu^2p^2 = N\delta^2Z^2_{1-\frac{\alpha}{2}} - n\delta^2Z^2_{1-\frac{\alpha}{2}}$$

$$n(N\mu^2p^2 + \delta^2Z^2_{1-\frac{\alpha}{2}}) = N\delta^2Z^2_{1-\frac{\alpha}{2}}$$

$$\boxed{n = \frac{N\delta^2}{\frac{N\mu^2p^2}{Z^2_{1-\frac{\alpha}{2}}} + \delta^2}}$$

ملاحظة:

- ✓ إذا كانت قيمة التباين δ^2 غير معلومة، نستخدم بدلها مقدرتها غير المتحيزة S^2.
- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

2.4. تقدير حجم العينة اللازم لتقدير المجموع الكلي للمجتمع T بمستوى دقة نسبي p'

يعطى مستوى الدقة النسبية لتقدير المجموع الكلي للمجتمع T بالعبارة:

$$p' = \left|\frac{\widehat{T} - T}{T}\right|$$

لدينا:

$$Z' = \frac{\widehat{T} - T}{\delta_{\widehat{T}}}$$

ومنها فإننا نجد:

$$Z'_{1-\frac{\alpha}{2}} = \frac{\left|\frac{\widehat{T} - T}{T}\right|}{\frac{\delta_{\widehat{T}}}{T}}$$

بتربيع الطرفين وتعويض ($T=N\mu$) يصبح لدينا:

$$Z'^2_{1-\frac{\alpha}{2}} = \frac{p'^2}{(1 - \frac{n}{N})\frac{\delta^2}{n\mu^2}}$$

وبنفس الطريقة كما في السابق نجد:

$$\boxed{n = \frac{N\delta^2}{\frac{N\mu^2 p'^2}{Z'^2_{1-\frac{\alpha}{2}}} + \delta^2}}$$

ملاحظة:

- ✓ إذا كانت قيمة التباين δ^2 غير معلومة، نستخدم بدلها مقدرتها غير المتحيزة S^2.
- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z' \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

مثال 3-4:

بالعودة لبيانات المثالين 3-1 و3-2، المطلوب:

1/ أوجد حجم العينة العشوائية البسيطة اللازم سحبها لتقدير متوسط الاستهلاك الشهري الأسري في هذه البلدية بحيث لا يتعدى الفرق المطلق النسبي بين قيمة μ الفعلية والمقدرة ال (0.04)، $p = \left|\frac{\bar{x}-\mu}{\mu}\right| = 0.04$، وذلك بمستوى ثقة 0.95 ؟

2/ أوجد حجم العينة العشوائية البسيطة اللازم سحبها لتقدير إجمالي الاستهلاك الشهري للأسر في هذه البلدية بحيث لا يتعدى الفرق المطلق النسبي بين قيمة T الفعلية والمقدرة ال $p' = \left|\frac{\hat{T}-T}{T}\right| = 0.04$، وذلك بمستوى ثقة 0.95 ؟

الحل:

1/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير μ بحيث لا يتعدى الفرق المطلق النسبي، $p = \left|\frac{\bar{x}-\mu}{\mu}\right| = 0.04$، وذلك بمستوى ثقة 0.95:

$$n = \frac{1000.(2)}{\frac{1000.(22)^2.(0{,}04)^2}{(1{,}96)^2} + 2} = 9{,}82$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 10$.

2/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير T بحيث لا يتعدى الفرق المطلق النسبي $p' = \left|\frac{\hat{T}-T}{T}\right| = 0.04$، وذلك بمستوى ثقة 0.95:

$$n = \frac{1000.(2)}{\frac{1000.(22)^2.(0{,}04)^2}{(1{,}96)^2} + 2} = 9{,}82$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 10$.

5. تقدير نسبة (P) وعدد (R) المفردات التي تتصف بصفة معينة في المجتمع

كثيرا ما يهتم الباحث بتقدير نسبة السكان اللذين يتصفون بصفة معينة، مثل نسبة الأمية بين الإناث، نسبة المدخنين بين طلبة الجامعة، معدل البطالة بين السكان النشطين إقتصاديا، نسبة من سيعطون أصواتهم لأحد المرشحين،... إلخ. حيث يلجأ الباحث إلى سحب عينة عشوائية لتقدير نسبة المفردات التي تتصف بالصفة المعينة P[17].

لتقدير نسبة صفة معينة P لمجتمع ما $(\hat{P})$، نضع: $P=\frac{R}{N}$، حيث R تمثل عدد الأفراد المتصفين بهذه الصفة المعينة من بين مجتمع حجمه N.

وللتعبير عن كون المفردة تتصف بالصفة المعينة من عدمها نستعمل المتغيرة X_i والتي تأخذ القيمتين 0 أو1، الواحد إذا كانت المفردة تتصف بالصفة المعينة والصفر إذا لم تكن تتصف بها. وعليه يصبح:

[17] فتحي أحمد عاروري، مرجع سبق ذكره، ص127.

$$\sum_{i=1}^{N} X_i = R$$

و

$$P = \frac{\sum_{i=1}^{N} X_i}{N} = \frac{R}{N}$$

وبالنسبة لعينة عشوائية بسيطة $x_1, x_2, \dots, x_n$، فإن عدد اللذين يتصفون بالصفة المعينة في هذه العينة هو r وتصبح النسبة السابقة:

$$\sum_{i=1}^{n} x_i = r$$

و

$$\widehat{P} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{r}{n}$$

مما سبق يتضح أن r لها توزيع ذو الحدين بمتوسط $E(r) = nP$ وتباين $V(r) = nPQ$ حيث Q=1-P، وأن $\widehat{P}$ تعتبر مقدرة غير متحيزة ل P لأن:

$$E(\widehat{P}) = E(\frac{r}{n}) = E(\frac{nP}{n}) = E(P) = P$$

بتباين معطى بالعلاقة: $V(\widehat{P}) = \hat{\delta}^2_{\widehat{P}} = \frac{\widehat{P}\widehat{Q}}{n-1}.(1 - \frac{n}{N})$ إذا كان السحب من دون إعادة وحجم العينة صغيرا $n \leq 30$.

أما إذا كانت $n > 30$، فتصبح $V(\widehat{P}) = \hat{\delta}^2_{\widehat{P}} = \frac{\widehat{P}\widehat{Q}}{n}.\left(1 - \frac{n}{N}\right)$.

و$V(\widehat{P}) = \hat{\delta}^2_{\widehat{P}} = \frac{\widehat{P}\widehat{Q}}{n}$ إذا كانت المعاينة من مجتمع غير محدود وبسحب مع الإعادة، حيث: $\widehat{Q} = 1-\widehat{P}$.

إذا كان حجم العينة كبيرا فإن: $\widehat{P} \sim N(P, \hat{\delta}^2_{\widehat{P}})$، وبالتالي: $Y = \frac{\widehat{P}-P}{\hat{\delta}_{\widehat{P}}} \sim N(0,1)$.

ومن خلال ما سبق، يمكن أيضا استخلاص مقدرة العدد الحقيقي R $(\widehat{R})$ للمفردات التي تتصف بهذه الصفة المعينة داخل مجتمع الدراسة، وفيما يلي ملخص لأهم خصائصها الاحصائية:

الجدول رقم 3-2: القيمة الفعلية والمقدرة للعدد R

	عدد المفردات المتصفة بهذه الصفة المعينة
القيمة الفعلية	$R = N.P$
القيمة المقدرة	$\widehat{R} = N.\widehat{P}$
المتوسط	$E(\widehat{R}) = N.P = R$
التباين الفعلي	$\delta_{\widehat{R}}^2 = N^2 \frac{PQ}{n}.\left(1 - \frac{n}{N}\right)$
التباين المقدر	$\hat{\delta}_{\widehat{R}}^2 = N^2 \frac{\widehat{P}\widehat{Q}}{n}.\left(1 - \frac{n}{N}\right)$
الانحراف المعياري	$\hat{\delta}_{\widehat{R}} = N\sqrt{\frac{\widehat{P}\widehat{Q}}{n}.\left(1 - \frac{n}{N}\right)}$

المصدر: من إعداد الأستاذ

6. مجال الثقة الخاص بالنسبة P وبعدد مفردات المجتمع المتصفة بالصفة المعينة R

1.6. مجال الثقة للنسبة P

مجال الثقة للنسبة P إذا كانت نسبة من يتصفون بصفة معينة في عينة حجمها n مأخوذة من مجتمع حجمه N هي $\widehat{P}$، وذلك عند مستوى ثقة $(1-\alpha)$ يمكن التوصل إليه كما يلي:

لدينا:

$$P(-y \le Y \le +y) = 1 - \alpha$$

أي:

$$-y \le Y \le +y$$

$$-y \le \frac{\widehat{P} - P}{\hat{\delta}_{\widehat{P}}} \le +y$$

$$\widehat{P} - y.\hat{\delta}_{\widehat{P}} \le P \le \widehat{P} + y.\hat{\delta}_{\widehat{P}}$$

$$P \in \left[\widehat{P} - y.\hat{\delta}_{\widehat{P}}\ ,\widehat{P} + y.\hat{\delta}_{\widehat{P}}\right]$$ أو

وعليه يكون مجال ثقة النسبة P عند مستوى ثقة (1-α) هو:

$$\boxed{P \in \left[\widehat{P} - y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{P}}\ ,\widehat{P} + y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{P}}\right]}$$

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا ($n \le 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

كما يمكن إحتساب مجال الثقة للنسبة P أيضا باستخدام طريقة حد الخطأ في تقدير النسبة P $(\widehat{P})$، حيث:

$$B^* = 2.\hat{\delta}_{\hat{P}}$$

أي أن نسبة من يتصفون بصفة معينة في مجتمع حجمه N (P)، وذلك عند مستوى ثقة $(1-\alpha)$ يقع في المدى (مجال الثقة) الآتي:

$$\boxed{P = \widehat{P} \mp B^*}$$

2.6. مجال الثقة للعدد الفعلي R

مجال الثقة للعدد الحقيقي R إذا كان عدد من يتصفون بصفة معينة في عينة حجمها n مأخوذة من مجتمع حجمه N هو $\widehat{R}$، وذلك عند درجة معنوية α يمكن التوصل إليه كما يلي:

لدينا: $Y' = \frac{\widehat{R}-R}{\hat{\delta}_{\widehat{R}}} \sim N(0,1)$

إذن: $P(-y' \leq Y' \leq +y') = 1-\alpha$

أي: $-y' \leq Y' \leq +y'$

$$-y' \leq \frac{\widehat{R}-R}{\hat{\delta}_{\widehat{R}}} \leq +y'$$

$$\widehat{R} - y'.\hat{\delta}_{\widehat{R}} \leq R \leq \widehat{R} + y'.\hat{\delta}_{\widehat{R}}$$

أو $R \in \left[\widehat{R} - y'.\hat{\delta}_{\widehat{R}}\ , \widehat{R} + y'.\hat{\delta}_{\widehat{R}}\right]$

وعليه يكون مجال ثقة R عند مستوى ثقة $(1-\alpha)$ هو:

$$\boxed{R \in \left[\widehat{R} - y'_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{R}}\ , \widehat{R} + y'_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{R}}\right]}$$

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا $(n \leq 30)$، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y' \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

كما يمكن إحتساب مجال الثقة ل R أيضا باستخدام طريقة حد الخطأ في تقدير العدد الحقيقي R $(\widehat{R})$، حيث:

$$D^* = 2.\hat{\delta}_{\widehat{R}}$$

أي أن عدد من يتصفون بصفة معينة في مجتمع حجمه N (R)، وذلك عند مستوى ثقة $(1-\alpha)$، يقع في المدى (مجال الثقة) الآتي:

$$\boxed{R = \widehat{R} \mp D^*}$$

مثال 3-5:

لدراسة نسبة العمال الذين يملكون سيارات شخصية بمؤسسة X والبالغ عددهم 2500 عامل، قررت الإدارة سحب عينة عشوائية بسيطة حجمها n=250 عامل لهذا الغرض، ومن خلالها تبين لنا أن r=100 عامل يملكون سيارة شخصية.

المطلوب:

1/ أوجد مجال ثقة نسبة عمال المؤسسة الذين يملكون سيارات شخصية P وذلك بدرجة معنوية 0.01.

2/ أوجد تقدير لعدد عمال المؤسسة الذين يملكون سيارات شخصية R وذلك بدرجة معنوية 0.01.

3/ إختبر الفرضية القائلة بأن نسبة عمال المؤسسة الذين يملكون سيارات شخصية P لا يزيد عن 0.33 وذلك بدرجة معنوية 0.01.

الحل:

1/ من خلال المعطيات نجد أن نسبة عمال المؤسسة الذين يملكون سيارات شخصية في العينة المختارة:

$$\widehat{P} = \frac{r}{n} = \frac{100}{250} = 0{,}4,\ \widehat{Q} = 1 - \widehat{P} = 1 - 0{,}4 = 0{,}6.$$

ويمكن حساب مقدرة تباين نسبة عمال المؤسسة الذين يملكون سيارات شخصية $\hat{\delta}^2_{\widehat{P}}$ حيث:

$$\hat{\delta}^2_{\widehat{P}} = \frac{\widehat{P}\widehat{Q}}{n}.\left(1 - \frac{n}{N}\right) = \frac{(0{,}4)(0{,}6)}{250}.\left(1 - \frac{250}{2500}\right) = 0{,}000864$$

الانحراف المعياري يساوي: $\hat{\delta}_{\widehat{P}} = \sqrt{0{,}000864} = 0{,}0293$.

من جدول التوزيع الطبيعي $N(0,1)$ نجد أن: $y_{1-\frac{\alpha}{2}} = 2{,}57$.

وعليه فإن مجال ثقة P وذلك عند مستوى ثقة 0.99 هو:

$$P \in [0{,}4 - (2{,}57).(0{,}0293)\ ,\ 0{,}4 + (2{,}57).(0{,}0293)]$$

$$P \in [0.324\ ,\ 0.475]$$

أي أن قيمة P لا يمكن أن تقل عن 0,324 ولا أن تزيد عن 0,475 وذلك عند مستوى ثقة 0.99.

2/ من خلال المعطيات نحتسب:

$$\widehat{R} = N\widehat{P} = 2500.(0{,}4) = 1000,$$

$$\hat{\delta}_{\hat{R}} = N.\hat{\delta}_{\hat{P}} = 2500.(0{,}0293) = 73{,}25.$$

وعليه فإن مجال ثقة R وذلك عند مستوى ثقة 0.99 هو:

$$R \in [1000 - (2{,}57).(73{,}25)\ ,\ 1000 + (2{,}57).(73{,}25)]$$

$$R \in [811{,}74\ ,\ 1188{,}25]$$

أي أن قيمة R لا يمكن أن تقل عن 812 عامل ولا أن تزيد عن 1189 عامل وذلك عند مستوى ثقة 0.99.

3/ الفرض القائل بأن نسبة عمال المؤسسة الذين يملكون سيارات شخصية P لا يزيد عن 0,33 وذلك بمستوى ثقة 0.99، يجعلنا أمام إختبار من طرف واحد على اليسار كالآتي:

$$\begin{cases} H_0 : P \leq 0{,}33 \\ H_1 : P > 0{,}33 \end{cases}$$

نقوم باحتساب الدالة الاختبارية $Z_{cal} = \frac{\hat{P}-P}{\hat{\delta}_{\hat{P}}} = \frac{0{,}4-0{,}33}{0{,}0293} = 2{,}38$.

نلاحظ أن قيمة Z_{cal} لا تنتمي لمجال منطقة قبول H_0: $]-\infty, 2.33]$[18]، إذن نرفض H_0 وبالتالي فإن نسبة عمال المؤسسة الذين يملكون سيارات شخصية P يزيد عن 0,33 وذلك عند مستوى ثقة 0.99.

7. تقدير حجم العينة اللازم لتقدير النسبة P والعدد الفعلي R

1.7. تقدير حجم العينة اللازم لتقدير النسبة P

بنفس الطريقة المعتمدة في تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ لدينا:

$$f = \left|\hat{P} - P\right|$$

$$Y = \frac{\left|\hat{P} - P\right|}{\hat{\delta}_{\hat{P}}}$$

بتربيع الطرفين يصبح لدينا:

$$Y^2_{1-\frac{\alpha}{2}} = \frac{f^2}{\frac{\hat{P}\hat{Q}}{n}.(1 - \frac{n}{N})}$$

وبنشر وجمع الحدود المتشابهة نجد:

$$\boxed{n = \frac{N\hat{P}\hat{Q}}{\frac{Nf^2}{Y^2_{1-\frac{\alpha}{2}}} + \hat{P}\hat{Q}}}$$

[18] من جدول التوزيع الطبيعي $N(0,1)$ في الملحق نجد أن: $z_{1-\alpha} = 1{,}65$.

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

يمكن أيضا حساب حجم العينة باستخدام طريقة حد الخطأ في تقدير النسبة P (B)، مثلما قمنا به في تقدير متوسط المجتمع μ فنجد:

$$n = \frac{N\widehat{P}\widehat{Q}}{ND + \widehat{P}\widehat{Q}}$$

حيث:

$B = 2.\delta_{\widehat{P}}, D = \frac{B^2}{4}, \widehat{Q} = 1 - \widehat{P}.$

تجدر الإشارة إلى أنه في حالة لم يتوفر تقدير أولي للنسبة P $(\widehat{P})$ فإنه يمكن استخدام $\widehat{P} = 0.5$ لتقدير حجم العينة.

2.7. تقدير حجم العينة اللازم لتقدير العدد الحقيقي R

بنفس الطريقة المعتمدة في تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ لدينا:

$$f' = \left|\widehat{R} - R\right|$$

$$Y' = \frac{\left|\widehat{R} - R\right|}{\widehat{\delta}_{\widehat{R}}}$$

بتربيع الطرفين يصبح لدينا:

$$Y'^2_{1-\frac{\alpha}{2}} = \frac{f'^2}{N^2.\frac{\widehat{P}\widehat{Q}}{n}.(1 - \frac{n}{N})}$$

وبنشر وجمع الحدود المتشابهة نجد:

$$n = \frac{N^2\widehat{P}\widehat{Q}}{\frac{f'^2}{Y'^2_{1-\frac{\alpha}{2}}} + N\widehat{P}\widehat{Q}}$$

ملاحظة:

إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y' \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

يمكن أيضا حساب حجم العينة باستخدام طريقة حد الخطأ في تقدير العدد الحقيقي R (B')، مثلما قمنا به في تقدير متوسط المجتمع μ فنجد:

$$n = \frac{N^2 \widehat{P}\widehat{Q}}{D' + N\widehat{P}\widehat{Q}}$$

حيث:

$B' = 2.\delta_{\widehat{R}}, D' = \frac{B'^2}{4}, \widehat{Q} = 1 - \widehat{P}.$

تجدر الإشارة إلى أنه في حالة لم يتوفر تقدير أولي للنسبة P $(\widehat{P})$ فإنه يمكن استخدام $\widehat{P} = 0.5$ لتقدير حجم العينة.

مثال 3-6:

بالعودة لبيانات المثال 3-5، المطلوب:

1/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير نسبة عمال المؤسسة الذين يملكون سيارات شخصية بحيث لا يتعدى الفرق المطلق بين قيمة P الحقيقة والمقدرة ال $f = |\widehat{P} - P| = 0.075$، وذلك بدرجة معنوية 0.01 ؟ من خلال النتيجة المتوصل إليها، ما الذي يجب على إدارة المؤسسة القيام به من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي ؟

2/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير عدد عمال المؤسسة الذين يملكون سيارات شخصية بحيث لا يتعدى الفرق المطلق بين قيمة R الحقيقة والمقدرة ال $f' = |\widehat{R} - R| = 185$، وذلك بدرجة معنوية 0.01 ؟ من خلال النتيجة المتوصل إليها، ما الذي يجب على إدارة المؤسسة القيام به من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي ؟

الحل:

1/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير P بحيث لا يتعدى الفرق المطلق، $f = |\widehat{P} - P| = 0.075$، وذلك بمستوى ثقة 0.99:

$$n = \frac{2500.(0{,}4).(0{,}6)}{\frac{2500.(0{,}075)^2}{(2{,}57)^2} + (0{,}4).(0{,}6)} = 253{,}26$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 254$. وبالتالي يجب على إدارة المؤسسة من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي إضافة 04 مفردات أخرى للعينة المختارة.

2/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير R بحيث لا يتعدى الفرق المطلق، $f' = |\widehat{R} - R| = 185$، وذلك بمستوى ثقة 0.99:

$$n = \frac{(2500)^2.(0,4).(0,6)}{\frac{(185)^2}{(2,57)^2} + 2500.(0,4).(0,6)} = 259,43$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 260$. وبالتالي يجب على إدارة المؤسسة من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي إضافة 10 مفردات أخرى للعينة المختارة.

8. تقدير حجم العينة اللازم لتقدير P وR باستخدام مستوى الدقة النسبي

يمكن أن نقوم أيضا بتقدير حجم العينة اللازم سحبها من أجل تحقيق مستوى معين من الدقة النسبية لتقدير النسبة P أو لتقدير عدد مفردات المجتمع التي تتصف بصفة معينة R.

1.8. تقدير حجم العينة اللازم لتقدير النسبة P بمستوى دقة نسبية h

ليكن:

$$h = \left|\frac{\widehat{P} - P}{P}\right|$$

$$Y = \frac{\left|\frac{\widehat{P} - P}{P}\right|}{\frac{\delta_{\widehat{P}}}{P}}$$

بتربيع الطرفين نجد:

$$Y^2_{1-\frac{\alpha}{2}} = \frac{h^2}{\frac{Q}{nP}.(1 - \frac{n}{N})}$$

$$h^2 = \frac{NQY^2_{1-\frac{\alpha}{2}} - nQY^2_{1-\frac{\alpha}{2}}}{nNP}$$

$$nNPh^2 = NQY^2_{1-\frac{\alpha}{2}} - nQY^2_{1-\frac{\alpha}{2}}$$

ومنه نجد:

$$\boxed{n = \frac{NQ}{\frac{NPh^2}{Y^2_{1-\frac{\alpha}{2}}} + Q}}$$

ملاحظة:

- ✓ في حالة عدم توفر القيم الفعلية P وQ نستعمل مقدراتها $(\widehat{P}, \widehat{Q})$.
- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y \sim t^{1-\frac{\alpha}{2}}_{n-1}$، بدل التوزيع الطبيعي $N(0,1)$.

2.8. تقدير حجم العينة اللازم لتقدير R بمستوى دقة نسبية h'

ليكن:

$$h' = \left|\frac{\widehat{R} - R}{R}\right|$$

$$Y' = \frac{\left|\frac{\widehat{R} - R}{R}\right|}{\frac{\delta_{\widehat{R}}}{R}}$$

بتربيع الطرفين نجد:

$$Y'^2_{1-\frac{\alpha}{2}} = \frac{h'^2}{\frac{Q}{nP}.(1 - \frac{n}{N})}$$

وبنشر وجمع الحدود المتشابهة نجد:

$$\boxed{n = \frac{NQ}{\frac{NPh'^2}{Y'^2_{1-\frac{\alpha}{2}}} + Q}}$$

ملاحظة:

- ✓ في حالة عدم توفر القيم الفعلية P وQ نستعمل مقدراتها $(\widehat{P}, \widehat{Q})$.
- ✓ إذا كان حجم العينة صغيرا $(n \leq 30)$، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y' \sim t^{1-\frac{\alpha}{2}}_{n-1}$، بدل التوزيع الطبيعي $N(0,1)$.

مثال 3-7:

بالعودة لبيانات المثال 3-5 وعلى اعتبار الدراسة بمثابة دراسة أولية، المطلوب:

1/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير نسبة عمال المؤسسة الذين يملكون سيارات شخصية بحيث لا يتعدى الفرق المطلق النسبي بين قيمة P الحقيقة والمقدرة ال $h = \left|\frac{\hat{P} - P}{P}\right| = 0.18$، وذلك بدرجة معنوية 0.01 ؟ من خلال النتيجة المتوصل إليها، ما الذي يجب على إدارة المؤسسة القيام به من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي ؟

2/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير عدد عمال المؤسسة الذين يملكون سيارات شخصية بحيث لا يتعدى الفرق المطلق النسبي بين قيمة R الحقيقة والمقدرة ال $h' = \left|\frac{\widehat{R} - R}{R}\right| = 0.18$، وذلك بدرجة معنوية 0.01 ؟ من خلال النتيجة المتوصل إليها، ما

الذي يجب على إدارة المؤسسة القيام به من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي ؟

الحل:

1/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير P بحيث لا يتعدى الفرق المطلق النسبي، $h = \left|\frac{\hat{P}-P}{P}\right| = 0.18$، وذلك بمستوى ثقة 0.99:

$$n = \frac{2500.(0,6)}{\frac{2500.(0,4).(0,18)^2}{(2,57)^2} + (0,6)} = 272,45$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 273$. وبالتالي يجب على إدارة المؤسسة من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي إضافة 23 مفردات أخرى للعينة المختارة.

2/ حجم العينة العشوائية البسيطة اللازم سحبها لتقدير R بحيث لا يتعدى الفرق المطلق النسبي، $h' = \left|\frac{\hat{R}-R}{R}\right| = 0.18$، وذلك بمستوى ثقة 0.99:

$$n = \frac{2500.(0,6)}{\frac{2500.(0,4).(0,18)^2}{(2,57)^2} + (0,6)} = 272,45$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 273$. وبالتالي يجب على إدارة المؤسسة من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي إضافة 23 مفردات أخرى للعينة المختارة.

9. تقدير حجم العينة الواجب سحبها لتقدير أكثر من صفة

لجمع البيانات المطلوبة عن الصفات المدروسة بالنسبة لمجتمع ما نلجأ في العادة إلى استخدام الاستمارة. ولتحديد حجم العينة اللازم لهذه الدراسة فإننا نقوم بذلك بالنسبة لكل صفة على حدى ثم نختار أكبر حجم عينة مقدر ليكون هو حجم العينة المطلوب، وهذا مع الأخذ بعين الاعتبار لعوامل التكلفة، الوقت، مستويي الدقة والمعنوية في السماح بالقيام بهذه دراسة.

مثال 3-8:

إذا أردنا تقدير الاستهلاك الشهري للأسرة الواحدة في أحد المناطق من مادتي الكهرباء والغاز لأغراض التدفئة بحيث لا يتعدى الفرق المطلق بين الاستهلاك الفعلي والمقدر من مادة الكهرباء ال (2) واط والفرق المطلق بين الاستهلاك الفعلي والمقدر من مادة الغاز ال (1.5) متر مكعب، وإذا علمنا بأن عدد الأسر في هذه المنطقة يقدر ب1.500 أسرة وأن تباين الاستهلاك الشهري للأسرة الواحدة من الكهرباء $\delta_1^2 = 12$ وتباين الاستهلاك الشهري للأسرة الواحدة من الغاز $\delta_2^2 = 8$.

المطلوب:

- ماهو حجم العينة العشوائية البسيطة اللازم سحبها لهذا الغرض وذلك عند درجة معنوية 0.05 ؟

الحل:

لتحديد حجم العينة العشوائية البسيطة اللازم سحبها في هذه الحالة نقوم باحتساب حجم العينة اللازم لتقدير الاستهلاك الشهري للأسرة الواحدة في هذه المنطقة بالنسبة لكل مادة على حدى، ثم نقوم باختيار أكبر حجم عينة بينهما.

-حجم العينة العشوائية البسيطة اللازم سحبها لتقدير الاستهلاك الشهري للأسرة الواحدة في هذه المنطقة من مادة الكهرباء μ بحيث لا يتعدى الفرق المطلق، $d = |\bar{x} - \mu| = 2$، وذلك بمستوى ثقة 0.95:

$$n = \frac{1500.(12)}{\frac{1500.(2)^2}{(1{,}96)^2} + (12)} = 11{,}43$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 12$.

-حجم العينة العشوائية البسيطة اللازم سحبها لتقدير الاستهلاك الشهري للأسرة الواحدة في هذه المنطقة من مادة الغاز μ بحيث لا يتعدى الفرق المطلق، $d = |\bar{x} - \mu| = 1{,}5$، وذلك بمستوى ثقة 0.95:

$$n = \frac{1500.(8)}{\frac{1500.(1{,}5)^2}{(1{,}96)^2} + (8)} = 13{,}53$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 14$.

مما سبق، يجب أن نختار عينة حجمها $n = 14$ أسرة على الأقل من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي.

سلسلة تمارين مقترحة حول المحور الثالث:

تمرين 01:

إذا كان لدينا مجتمع محدود حجمه N=100 وقررنا سحب عينة عشوائية بسيطة حجمها n=5.

المطلوب:

-ماهو عدد العينات التي يمكن سحبها من هذا المجتمع ؟

تمرين 02:

إذا كان لدينا مجتمع محدود مفرداته $\{1, 2, 3\}$.

المطلوب:

1/ أحسب متوسط المجتمع μ وإنحرافه المعياري δ.

2/ إذا أردنا سحب عينات حجمها (n=2)، من خلال السحب بإعادة والسحب من دون إعادة.

بالنسبة للسحب بإعادة:

- ماهو عدد العينات الممكن سحبها من هذا المجتمع ؟
- أذكر هذه العينات بالتفصيل.
- أحسب متوسط كل عينة.
- أحسب متوسط متوسطات العينات المسحوبة $\mu_{\bar{x}}$.
- أحسب الإنحراف المعياري لمتوسطات العينات المسحوبة $\delta_{\bar{x}}$.
- هل يمكن تأكيد العلاقتين الآتيتين: $\delta_{\bar{x}} = \frac{\delta}{\sqrt{n}}$, $\mu_{\bar{x}} = \mu$.

بالنسبة للسحب بدون إعادة:

- ماهو عدد العينات الممكن سحبها من هذا المجتمع ؟
- أذكر هذه العينات بالتفصيل.
- أحسب متوسط كل عينة.
- أحسب متوسط متوسطات العينات المسحوبة $\mu_{\bar{x}}$.
- أحسب الإنحراف المعياري لمتوسطات العينات المسحوبة $\delta_{\bar{x}}$.
- هل يمكن تأكيد العلاقتين الآتيتين: $\delta_{\bar{x}} = \frac{\delta}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N-1}}$, $\mu_{\bar{x}} = \mu$.

تمرين 03:

إذا كانت نقاط 400 طالب في إمتحان مادة تقنيات المعاينة تتبع التوزيع الطبيعي بمتوسط يساوي 9,9 وإنحراف معياري قدره 3,57، فماهو إحتمال أن يكون متوسط نقاط عينة عشوائية بسيطة مكونة من 50 طالب:

1/ محصور بين 10 و13,5.

2/ أقل من 10.

ملاحظة: أوجد الإحتمال في حالتي السحب بإعادة وبدون إعادة.

تمرين 04:

لدراسة الاستهلاك الشهري لغاز المدينة للأسرة الواحدة في إحدى البلديات، سحبنا عينة عشوائية بسيطة حجمها n=25 أسرة، فكان الوسط الحسابي للإستهلاك هو $\bar{x}$ =22 متر مكعب وتباين $s^2 = 2$.

المطلوب:

1/ أوجد مجال الثقة بالنسبة لمتوسط الاستهلاك الشهري للأسرة الواحدة μ عند مستوى ثقة 0.99.

2/ إختبر الفرضية القائلة بأن متوسط الاستهلاك الشهري للأسرة الواحدة في هذه البلدية لا يقل عن 22.5 متر مكعب وذلك بمستوى ثقة 0.99.

3/ إذا كانت البلدية المعنية مشكلة من 250 أسرة، فماهو حجم الاستهلاك الشهري لغاز المدينة المتوقع في هذه البلدية وذلك بمستوى ثقة 0.99 ؟

تمرين 05:

لأجل دراسة متوسط الزيادة الشهرية في النفقات الاستهلاكية بالنسبة للأسرة الواحدة خلال شهر رمضان بالمقارنة مع بقية الأشهر الأخرى بإحدى المناطق المكونة من 2500 أسرة، قمنا بسحب عينة عشوائية بسيطة حجمها n=100 أسرة لهذا الغرض من خلال دراسة النفقات الاستهلاكية الاضافية خلال آخر شهر رمضان قبل العملية وكانت النتائج كما يلي (الوحدة: دينار جزائري- دج):

النفقات الاستهلاكية الاضافية الشهرية	عدد الأسر
2000-0	10
4000-2000	35
6000-4000	20
8000-6000	15
10000-8000	15
12000-10000	5
المجموع	100

المطلوب:

1/ أوجد مجال الثقة بالنسبة لمتوسط النفقات الاستهلاكية الاضافية الشهرية للأسرة الواحدة خلال شهر رمضان في هذه المنطقة μ عند مستوى ثقة 0.95.

2/ إختبر الفرضية القائلة بأن متوسط النفقات الاستهلاكية الاضافية الشهرية للأسرة الواحدة خلال شهر رمضان في هذه المنطقة μ لا يزيد عن 4500 دج وذلك بمستوى ثقة 0.95.

3/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير متوسط النفقات الاستهلاكية الاضافية الشهرية للأسرة الواحدة خلال شهر رمضان بحيث لا يتعدى الفرق المطلق بين متوسط النفقات الاستهلاكية الاضافية الفعلية μ والمقدرة $\bar{x}$ ال 510 دج، $d = |\bar{x} - \mu| =$ 510، وذلك بمستوى ثقة 0.95 ؟

4/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير متوسط النفقات الاستهلاكية الاضافية الشهرية للأسرة الواحدة خلال شهر رمضان بحيث لا يتعدى الفرق النسبي المطلق بين متوسط النفقات الاستهلاكية الاضافية الفعلية والمقدرة ال (0.1)، $p = \left|\frac{\bar{x}-\mu}{\mu}\right| = 0.1$، وذلك بمستوى ثقة 0.95 ؟

5/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها بحيث يكون الحد على الخطأ في تقدير متوسط النفقات الاستهلاكية الاضافية الشهرية B=510 ؟

تمرين 06:

بالاعتماد على بيانات التمرين رقم 05 وعلى اعتبار هذه الدراسة بمثابة دراسة أولية:

1/ أوجد مجال الثقة بالنسبة لإجمالي النفقات الاستهلاكية الاضافية الشهرية T خلال شهر رمضان في هذه المنطقة وذلك عند مستوى ثقة 0.95.

2/ إختبر الفرضية القائلة بأن إجمالي النفقات الاستهلاكية الاضافية الشهرية T خلال شهر رمضان في هذه المنطقة لا يقل عن 14.000.000 دج وذلك بمستوى ثقة 0.95.

3/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير اجمالي النفقات الاستهلاكية الاضافية الشهرية T خلال شهر رمضان في هذه المنطقة بحيث لا يتعدى الفرق المطلق بين المجموع الكلي الفعلي T والمقدر $\widehat{T}$ ال 510.000 دج، $d' = \left|\widehat{T} - T\right| = 510.000$، وذلك بمستوى ثقة 0.95 ؟

4/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير اجمالي النفقات الاستهلاكية الاضافية الشهرية T خلال شهر رمضان في هذه المنطقة بحيث لا يتعدى الفرق النسبي المطلق بين المجموع الكلي الفعلي T والمقدر $\widehat{T}$ ال $p' = \left|\frac{\widehat{T}-T}{T}\right| = 0.04$، وذلك بمستوى ثقة 0.95؟

5/ ماهو حجم العينة اللازم سحبها لتقدير اجمالي النفقات الاستهلاكية الاضافية الشهرية T خلال شهر رمضان في هذه المنطقة بحيث يكون الحد على الخطأ في التقدير $B' = 1.275.000$ ؟

تمرين 07:

إذا كانت نسبة المسامير غير السليمة المنتجة من طرف مصنع ما مقدرة ب 3%، وقمنا بأخذ عينة حجمها n=300 مفردة، المطلوب:

- ما هو إحتمال ان تكون نسبة المسامير غير السليمة في هذه العينة تتراوح بين 2% و3,5%؟

تمرين 08:

لدراسة نسبة المدخنين بين طلبة كلية الطب سحبنا عينة عشوائية من n=400 طالب، ومن خلالها تبين لنا أن r=200 طالب منهم يدخنون، فإذا علمنا بأن عدد طلبة كلية الطب هو N=5.000 طالب، **المطلوب:**

1/ أوجد تقدير لنسبة المدخنين بين طلبة كلية الطب وذلك بدرجة معنوية 0.01.

2/ إختبر الفرضية القائلة بأن نسبة المدخنين بين طلبة كلية الطب لا يقل عن 0.4 وذلك بدرجة معنوية 0.01.

3/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير نسبة المدخنين بين طلبة كلية الطب بحيث لا يتعدى الفرق المطلق بين النسبة الحقيقة P والنسبة المقدرة $\widehat{P}$ ال $f = \left|\widehat{P} - P\right| =$ 0.04، وذلك بدرجة معنوية 0.01 ؟

4/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير نسبة المدخنين بين طلبة كلية الطب بحيث لا يتعدى الفرق النسبي المطلق بين النسبة الحقيقة P والنسبة المقدرة $\widehat{P}$ ال $h = \left|\frac{\widehat{P}-P}{P}\right| = 0.08$، وذلك بدرجة معنوية 0.01 ؟

5/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير نسبة المدخنين بين طلبة كلية الطب بحيث يكون الحد على الخطأ في التقدير $B = 0.04$ ؟

<u>تمرين 09:</u>

بالاعتماد على بيانات التمرين رقم 08 وعلى اعتبار هذه الدراسة بمثابة دراسة أولية:

1/ أوجد تقدير لعدد المدخنين بين طلبة كلية الطب وذلك بمستوى ثقة 0.99.

2/ إختبر الفرضية القائلة بأن عدد المدخنين بين طلبة كلية الطب لا يزيد عن 2.200 طالب وذلك بمستوى ثقة 0.99.

3/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير عدد المدخنين بين طلبة كلية الطب بحيث لا يتعدى الفرق المطلق f' بين العدد الحقيقي والمقدر ال 200 طالب وذلك بمستوى ثقة 0.99 ؟

4/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها لتقدير عدد المدخنين بين طلبة كلية الطب بحيث لا يتعدى الفرق النسبي المطلق h' بين العدد الحقيقي والمقدر للطلبة ال 0.08 وذلك بمستوى ثقة 0.99 ؟

5/ ماهو حجم العينة العشوائية البسيطة اللازم سحبها بحيث يكون الحد على الخطأ في تقدير عدد المدخنين بين طلبة كلية الطب B'=200 طالب ؟

<u>تمرين 10:</u>

إذا أردنا تقدير الاستهلاك السنوي للأسرة الواحدة في أحد الأرياف من مادتي زيت المائدة وزيت الزيتون بحيث لا يتعدى الفرق المطلق بين الاستهلاك الفعلي والمقدر من مادة زيت المائدة ال (1.5) لتر والفرق المطلق بين الاستهلاك الفعلي والمقدر من مادة زيت الزيتون ال

(1) لتر، وإذا علمنا بأن عدد الأسر في هذه المنطقة يقدر ب1.000 أسرة وأن تباين الاستهلاك السنوي للأسرة الواحدة من زيت المائدة $\delta_1^2 = 5$ وتباين الاستهلاك السنوي للأسرة الواحدة من زيت الزيتون $\delta_2^2 = 4$.

المطلوب:

ماهو حجم العينة العشوائية البسيطة اللازم سحبها لهذا الغرض وذلك عند درجة معنوية 0.01 ؟

المحور الرابع: العينة العشوائية الطبقية

لسحب العينة العشوائية الطبقية التي حجمها n من مجتمع حجمه N، نقوم بسحب عينات جزئية n_i من الطبقات المخصصة لها والتي حجمها N_i وذلك باستخدام المعاينة العشوائية البسيطة، وفي هذه الحالة تعتبر كل طبقة من المجتمع وكأنها مجتمع مستقل بذاته عن بقية الطبقات. وعليه يكون:

$$N = N_1+N_2+ \dots +N_k = \sum_{i=1}^{k} N_i$$

$$n = n_1+n_2+ \dots +n_k = \sum_{i=1}^{k} n_i$$

حيث يمثل k عدد طبقات المجتمع.

كما أن حجم العينة العشوائية البسيطة المسحوبة من الطبقة (i) يحدد بالعلاقة التالية:

$$n_i = n \cdot \frac{N_i}{N} = n \cdot W_i$$

حيث: $W_i = \frac{N_i}{N}$ يمثل وزن الطبقة i في المجتمع.

1. تقدير متوسط المجتمع μ والمجموع الكليT

على إعتبار x_{ij} تمثل قيمة الوحدة j في الطبقة i، فإن الوسط الحسابي للطبقة $\bar{x}_i$ يعطى بالعلاقة:

$$\bar{x}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} x_{ij}$$

وتباين s_i^2 الذي يعتبر مقدرة غير متحيزة لتباين الطبقة δ_i^2 ويعطى بالعلاقة:

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (x_{ij} - \bar{x}_i)^2$$

في حالة المعاينة من مجتمع محدود حجمه N وقيمة δ^2 مجهولة وبسحب من دون إعادة فإن القيم الحقيقية والمقدرة لكل من μ وT تكون (أنظر الجدول رقم 4-1):

الجدول رقم 4-1: القيم الفعلية والمقدرة للمتوسط μ والمجموع الكلي T

	المتوسط	المجموع الكلي
الطبقة (المعلمة الحقيقية)	$\mu_i = \frac{1}{N_i}\sum_{j=1}^{N_i} X_{ij}$	$T_i = \sum_{j=1}^{N_i} X_{ij} = N_i\mu_i$
الطبقة (المقدرة)	$\bar{x}_i = \frac{1}{n_i}\sum_{j=1}^{n_i} x_{ij}$	$\hat{T}_i = N_i\bar{x}_i$
تباين الطبقة	$\delta^2_{\bar{x}_i} = (1-\frac{n_i}{N_i})\frac{\delta_i^2}{n_i}$	$\delta^2_{\hat{T}_i} = N_i^2\delta^2_{\bar{x}_i}$
التباين المقدر	$\hat{\delta}^2_{\bar{x}_i} = (1-\frac{n_i}{N_i})\frac{s_i^2}{n_i}$	$\hat{\delta}^2_{\hat{T}_i} = N_i^2\hat{\delta}^2_{\bar{x}_i}$
الانحراف المعياري	$\hat{\delta}_{\bar{x}_i} = \sqrt{(1-\frac{n_i}{N_i})\frac{s_i^2}{n_i}}$	$\hat{\delta}_{\hat{T}_i} = N_i\hat{\delta}_{\bar{x}_i}$
المجتمع (المعلمة الحقيقية)	$\mu = \sum_{i=1}^{k}\frac{N_i}{N}\mu_i = \sum_{i=1}^{k} W_i\mu_i$	$T = \sum_{i=1}^{k} T_i = \sum_{i=1}^{k} N_i\mu_i = N\mu$
المجتمع (المقدرة)	$\bar{x} = \sum_{i=1}^{k}\frac{N_i}{N}\bar{x}_i = \sum_{i=1}^{k} W_i\bar{x}_i$	$\hat{T} = \sum_{i=1}^{k}\hat{T}_i = \sum_{i=1}^{k} N_i\bar{x}_i = N\bar{x}$
تباين المجتمع	$\delta^2_{\bar{x}} = \sum_{i=1}^{k} W_i^2\delta^2_{\bar{x}_i}$	$\delta^2_{\hat{T}} = N^2\delta^2_{\bar{x}}$
التباين المقدر	$\hat{\delta}^2_{\bar{x}} = \sum_{i=1}^{k} W_i^2\hat{\delta}^2_{\bar{x}_i}$	$\hat{\delta}^2_{\hat{T}} = N^2\hat{\delta}^2_{\bar{x}}$
الانحراف المعياري	$\hat{\delta}_{\bar{x}} = \sqrt{\sum_{i=1}^{k} W_i^2\hat{\delta}^2_{\bar{x}_i}}$	$\hat{\delta}_{\hat{T}} = N\hat{\delta}_{\bar{x}}$

المصدر: من إعداد الأستاذ

2. مجال الثقة بالنسبة لمتوسط المجتمع μ والمجموع الكلي T

1.2. مجال الثقة لمتوسط المجتمع μ

كما رأينا سابقا بالنسبة للمعاينة العشوائية البسيطة، فإن مجال الثقة بالنسبة لمتوسط المجتمع μ غير المعلوم بمستوى معنوية α يمكن حسابه كما يلي:

بما أن $\bar{x}$ هو متغير عشوائي يتبع التوزيع الطبيعي بمتوسط μوتباين $\delta_{\bar{x}}^2$، ومن خلال نظرية النهاية المركزية (TCL) فإن:

$Z = \frac{\bar{x}-\mu}{\delta_{\bar{x}}}$ يتبع التوزيع الاحتمالي الطبيعي بمتوسط معدوم وتباين يساوي الواحد $Z \sim N(0,1)$ عندما يكون حجم العينة كبيرا (n>30).

مع العلم أن:

$$P(-z \leq Z \leq +z) = 1 - \alpha$$

أي:

$$-z \leq Z \leq +z$$

$$-z \leq \frac{\bar{x} - \mu}{\delta_{\bar{x}}} \leq +z$$

وعليه يكون مجال الثقة بالنسبة ل μ بمستوى ثقة (1-α) هو:

$$\mu \in \left[\bar{x} - z_{1-\frac{\alpha}{2}}.\delta_{\bar{x}} , \bar{x} + z_{1-\frac{\alpha}{2}}.\delta_{\bar{x}}\right]$$

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا (n≤30)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

2.2. مجال الثقة للمجموع الكلي للمجتمع T

كما رأينا سابقا بالنسبة لمجال ثقة متوسط المجتمع μ، فإن مجال الثقة بالنسبة للمجموع الكلي للمجتمع T غير المعلوم بمستوى معنوية α يمكن حسابه كما يلي:

بما أن تقدير المجموع الكلي للمجتمع $\hat{T}$ هو متغير عشوائي يتبع التوزيع الطبيعي بمتوسط T وتباين $\delta_{\hat{T}}^2$، ومن خلال نظرية النهاية المركزية (TCL) فإن:

$Z' = \frac{\hat{T}-T}{\delta_{\hat{T}}}$ يتبع التوزيع الاحتمالي الطبيعي بمتوسط معدوم وتباين يساوي الواحد $Z' \sim N(0,1)$ عندما يكون حجم العينة كبيرا (n>30).

مع العلم أن:

$$P(-z' \leq Z' \leq +z') = 1 - \alpha$$

أي:

$$-z' \leq Z' \leq +z'$$

$$-z' \leq \frac{\hat{T} - T}{\delta_{\hat{T}}} \leq +z'$$

وعليه يكون مجال الثقة بالنسبة ل T بمستوى ثقة (1-α) هو:

$$T \in \left[\hat{T} - z'_{1-\frac{\alpha}{2}}.\delta_{\hat{T}} , \hat{T} + z'_{1-\frac{\alpha}{2}}.\delta_{\hat{T}}\right]$$

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z' \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

مثال 4-1:

لدراسة متوسط وزن المواليد خلال الشهر الأول من السنة الحالية بأحد المستشفيات، وعلى اعتبار المواليد موزعين على طبقتين على التوالي (ذكور وإناث) تقرر الاعتماد على المعاينة العشوائية الطبقية لأجل هذا الغرض. وكانت النتائج كالآتي:

الطبقة i	حجم الطبقة N_i	متوسط وزن العينة/ كغ $\bar{x}_i$	تباين العينة s_i^2
1	362	4,5	0,09
2	238	3,9	0,07
المجموع	600		

المطلوب:

1/ أوجد مجال الثقة لمتوسط وزن المواليد خلال الشهر الأول في هذا المستشفى μ عند درجة معنوية 5%.

2/ أوجد مجال الثقة لإجمالي أوزان المواليد خلال الشهر الأول في هذا المستشفى T عند درجة معنوية 5%.

الحل:

1/ من خلال المعطيات نقوم بإحتساب:

$$\bar{x} = \sum_{i=1}^{2} \frac{N_i}{N} \bar{x}_i = \frac{362}{600}.(4,5) + \frac{238}{600}.(3,9) = 4,262.$$

$$\hat{\delta}_{\bar{x}}^2 = \sum_{i=1}^{2} W_i^2 \hat{\delta}_{\bar{x}_i}^2 = \left(\frac{362}{600}\right)^2.(0,09) + \left(\frac{238}{600}\right)^2.(0,07) = 0,043.$$

$$\hat{\delta}_{\bar{x}} = \sqrt{0,043} = 0,209.$$

من جدول التوزيع الطبيعي $N(0,1)$ نجد أن: $z_{1-\frac{\alpha}{2}} = 1,96$.

وعليه فإن مجال ثقة μ وذلك عند مستوى ثقة 0.95 هو:

$$\mu \in [4,262 - (1,96).(0,209)\ ,\ 4,262 + (1,96).(0,209)]$$

$$\mu \in [3.852\ ,\ 4.671]$$

أي أن قيمة μ لا يمكن أن تقل عن 3,852كغ ولا أن تزيد عن 4,671كغ وذلك عند مستوى ثقة 0.95.

2/ لاحتساب مجال ثقة T:

$$\widehat{T} = N\bar{x} = 600.(4{,}262) = 2.557{,}2$$
$$\hat{\delta}_{\widehat{T}} = N\hat{\delta}_{\bar{x}} = 600.(0{,}209) = 125{,}4$$

وعليه فإن مجال ثقة T وذلك عند مستوى ثقة 0.95 هو:

$$T \in [2557{,}2 - (1{,}96).(125{,}4)\ ,\ 2557{,}2 + (1{,}96).(125{,}4)]$$
$$T \in [2311.41\ ,\ 2802.98]$$

أي أن قيمة T لا يمكن أن تقل عن 2311.41 كغ ولا أن تزيد عن 2802.98 كغ وذلك عند مستوى ثقة 0.95.

3. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ والمجموع الكلي للمجتمع T

إن لحجم العينة دورا هاما في تقدير μ وT وكذا تعميم النتائج المتوصل إليها. فمن المعلوم أنه كلما زاد حجم العينة كلما كانت هذه النتائج أكثر دقة وقريبة أكثر من معالم المجتمع محل الدراسة. غير أن أخذ عينة كبيرة قد لا يكون ضمن الامكانات المتاحة لذا نلجأ لتقدير حجم العينة اللازم وفق معايير محددة مسبقا.

1.3. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ بمستوى دقة d

يمكن تقدير حجم العينة العشوائية الطبقية n باستخدام مستوى دقة مناسب، حيث لا يزيد الفرق المطلق بين متوسط العينة ومتوسط المجتمع عن $d = |\bar{x} - \mu|$.

علما أن:

$$Z = \frac{\bar{x} - \mu}{\delta_{\bar{x}}} \sim N(0\ ,1)$$

فإننا نجد:

$$Z_{1-\frac{\alpha}{2}} = \frac{|\bar{x} - \mu|}{\delta_{\bar{x}}}$$

$$Z^2_{1-\frac{\alpha}{2}} = \frac{d^2}{\delta^2_{\bar{x}}}$$

$$Z^2_{1-\frac{\alpha}{2}} = \frac{d^2}{\sum_{i=1}^{k} W_i^2 \delta^2_{\bar{x}_i}} = \frac{d^2}{\sum_{i=1}^{k} W_i^2 (1 - \frac{n_i}{N_i}) \frac{\delta_i^2}{n_i}}$$

$$d^2 = Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i^2.\frac{\delta_i^2}{n_i} - Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{N}$$

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة على طبقات المجتمع $(n_i = \frac{n}{k})$ فإن:

$$d^2 = k.Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i^2.\frac{\delta_i^2}{n} - Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{N}$$

$$\boxed{n = \frac{k.\sum_{i=1}^{k} N_i^2.\delta_i^2}{\frac{N^2d^2}{Z^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة على طبقات المجتمع $(n_i = n\,.\,W_i)$ فإن:

$$d^2 = Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{n} - Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{N}$$

$$\boxed{n = \frac{N.\sum_{i=1}^{k} N_i.\delta_i^2}{\frac{N^2d^2}{Z^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة على طبقات المجتمع $(n_i = n\,.\,\frac{\frac{N_i\delta_i}{\sqrt{c_i}}}{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}})$، أين تمثل c_i تكلفة سحب المفردة الواحدة من الطبقة i، فإن:[19]

$$N^2d^2 = Z^2_{1-\frac{\alpha}{2}}.\frac{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}.\sum_{i=1}^{k} N_i\ddot{a}_i\sqrt{c_i}}{n} - Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} N_i.\delta_i^2$$

$$\boxed{n = \frac{(\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}).(\sum_{i=1}^{k} N_i\delta_i\sqrt{c_i})}{\frac{N^2d^2}{Z^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

ملاحظة:

- ✓ إذا كانت قيم تباين الطبقات δ_i^2 غير معلومة، نستخدم بدلها مقدراتها غير المتحيزة s_i^2.
- ✓ إذا كان حجم العينة صغيرا $(n \leq 30)$، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z \sim t^{1-\frac{\alpha}{2}}_{n-1}$، بدل التوزيع الطبيعي $N(0,1)$.

[19] Pascal ARDILLY, "Les techniques de sondage", Édition TECHNIP, Paris 2006, Page 96.

مثال 4-2:

بالعودة لبيانات المثال 4-1، **المطلوب:**

- تقدير حجم العينة العشوائية الطبقية اللازم سحبها لتقدير μ وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق بين المتوسط الفعلي والمقدر، $d = |\bar{x} - \mu| = 0,1$ وذلك بدرجة معنوية 5%.

ملاحظة: تكاليف سحب المفردة الواحدة من كل طبقة هي كالآتي:

$c_1 = 1.20, c_2 = 1.25.$

الحل:

-حجم العينة العشوائية الطبقية اللازم سحبها لتقدير μ وذلك باستخدام أساليب التخصيص المختلفة:

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة نجد:

$$n = \frac{2.(15759,04)}{\frac{(600)^2.(0,1)^2}{(1,96)^2} + 49,24} = 31,95$$

وعليه فإن حجم العينة اللازم سحبها هو $n = 32$ مولود جديد من بينهم 19 ذكر و13 أنثى (بالنظر لوزن كل طبقة في المجتمع الكلي W_i).

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة نجد:

$$n = \frac{600.(49,24)}{\frac{(600)^2.(0,1)^2}{(1,96)^2} + 49,24} = 29,95$$

وعليه فإن حجم العينة اللازم سحبها هو $n = 30$ مولود جديد من بينهم 18 ذكر و12 أنثى (بالنظر لوزن كل طبقة في المجتمع الكلي W_i).

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة نجد:

$$n = \frac{(155,45).(189,36)}{\frac{(600)^2.(0,1)^2}{(1,96)^2} + 49,24} = 29,85$$

وعليه فإن حجم العينة اللازم سحبها هو $n = 30$ مولود جديد من بينهم 18 ذكر و12 أنثى (بالنظر لوزن كل طبقة في المجتمع الكلي W_i).

2.3. تقدير حجم العينة اللازم لتقدير المجموع الكلي T بمستوى دقة d'

مما سبق وبنفس الطريقة المعتمدة في تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ لدينا:

$$\hat{T} \sim N(T, \delta_{\hat{T}})$$

$$\frac{\hat{T} - T}{\delta_{\hat{T}}} \sim N(0, 1)$$

ومنها فإننا نجد:

$$Z' = \frac{\left|\widehat{T} - T\right|}{\delta_{\widehat{T}}}$$

بوضع $d' = \left|\widehat{T} - T\right|$ وتربيع الطرفين نجد:

$$Z'^2_{1-\frac{\alpha}{2}} = \frac{d'^2}{\delta^2_{\widehat{T}}}$$

بتعويض $\delta^2_{\widehat{T}}$ بعبارته نجد:

$$Z'^2_{1-\frac{\alpha}{2}} = \frac{d'^2}{N^2 \sum_{i=1}^{k} W_i^2 \frac{\delta_i^2}{n_i} - N^2 \sum_{i=1}^{k} W_i \frac{\delta_i^2}{N}}$$

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة على طبقات المجتمع $(n_i = \frac{n}{k})$ فإن:

$$d'^2 = k.N^2.Z'^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i^2.\frac{\delta_i^2}{n} - N^2.Z'^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{N}$$

$$\boxed{n = \frac{k.\sum_{i=1}^{k} N_i^2.\delta_i^2}{\frac{d'^2}{Z'^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة على طبقات المجتمع $(n_i = n\,.\,W_i)$ فإن:

$$d'^2 = N^2.Z'^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i\frac{\delta_i^2}{n} - N^2.Z'^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i\frac{\delta_i^2}{N}$$

$$\boxed{n = \frac{N.\sum_{i=1}^{k} N_i.\delta_i^2}{\frac{d'^2}{Z'^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة على طبقات المجتمع $(n_i = n\,.\,\frac{\frac{N_i\delta_i}{\sqrt{c_i}}}{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}})$ فإن:

$$d'^2 = Z'^2_{1-\frac{\alpha}{2}}.\frac{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}.\sum_{i=1}^{k} N_i\delta_i\sqrt{c_i}}{n} - Z'^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} N_i.\delta_i^2$$

$$n = \frac{(\sum_{i=1}^{k} \frac{N_i \delta_i}{\sqrt{c_i}}) \cdot (\sum_{i=1}^{k} N_i \delta_i \sqrt{c_i})}{\frac{d'^2}{Z'^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i . \delta_i^2}$$

ملاحظة:

- ✓ إذا كانت قيم تباين الطبقات δ_i^2 غير معلومة، نستخدم بدلها مقدراتها غير المتحيزة s_i^2.
- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z' \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

4. تقدير حجم العينة اللازم لتقدير μ وT باستخدام مستوى دقة النسبي

يمكن أن نقوم أيضا بتقدير حجم العينة اللازم سحبها من أجل تحقيق مستوى معين من الدقة النسبية في التقدير.

1.4. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ بمستوى دقة نسبي p

يمكن أن نقوم أيضا بتقدير حجم العينة اللازم سحبها من أجل تحقيق مستوى معين من الدقة النسبية p لتقدير متوسط المجتمع μ.

يعطى مستوى الدقة النسبية لتقدير متوسط المجتمع μ بالعبارة:

$$p = \left| \frac{\bar{x} - \mu}{\mu} \right|$$

لدينا:

$$Z = \frac{\bar{x} - \mu}{\delta_{\bar{x}}}$$

ومنها فإننا نجد:

$$Z_{1-\frac{\alpha}{2}} = \frac{\left| \frac{\bar{x} - \mu}{\mu} \right|}{\frac{\delta_{\bar{x}}}{\mu}}$$

بتربيع الطرفين يصبح لدينا:

$$Z^2_{1-\frac{\alpha}{2}} = \frac{p^2}{\sum_{i=1}^{k} W_i^2 (1 - \frac{n_i}{N_i}) \frac{\delta_i^2}{n_i \mu^2}}$$

$$p^2 = Z^2_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i^2 . \frac{\delta_i^2}{n_i \mu^2} - Z^2_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i . \frac{\delta_i^2}{N \mu^2}$$

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة على طبقات المجتمع $(n_i = \frac{n}{k})$ فإن:

$$p^2 = k.Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i^2.\frac{\delta_i^2}{n\mu^2} - Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{N\mu^2}$$

$$\boxed{n = \frac{k.\sum_{i=1}^{k} N_i^2.\delta_i^2}{\frac{N^2\mu^2p^2}{Z^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة على طبقات المجتمع $(n_i = n\,.\,W_i)$ فإن:

$$p^2 = Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{n\mu^2} - Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{\delta_i^2}{N\mu^2}$$

$$\boxed{n = \frac{N.\sum_{i=1}^{k} N_i.\delta_i^2}{\frac{N^2\mu^2p^2}{Z^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة على طبقات المجتمع $(n_i = n\,.\,\frac{\frac{N_i\delta_i}{\sqrt{c_i}}}{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}})$ فإن:

$$N^2\mu^2p^2 = Z^2_{1-\frac{\alpha}{2}}.\frac{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}.\sum_{i=1}^{k} N_i\delta_i\sqrt{c_i}}{n} - Z^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} N_i.\delta_i^2$$

$$\boxed{n = \frac{(\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}).(\sum_{i=1}^{k} N_i\delta_i\sqrt{c_i})}{\frac{N^2\mu^2p^2}{Z^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}}$$

ملاحظة:

- ✓ إذا كانت قيم تباين الطبقات δ_i^2 غير معلومة، نستخدم بدلها مقدراتها غير المتحيزة s_i^2.
- ✓ إذا كان حجم العينة صغيرا $(n \leq 30)$، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z \sim t^{1-\frac{\alpha}{2}}_{n-1}$، بدل التوزيع الطبيعي $N(0,1)$.

2.4. تقدير حجم العينة اللازم لتقدير المجموع الكلي T بمستوى دقة نسبي p'

يعطى مستوى الدقة النسبية لتقدير T بالعبارة:

$$p' = \left|\frac{\widehat{T} - T}{T}\right|$$

لدينا:

$$Z' = \frac{\left|\widehat{T} - T\right|}{\delta_{\widehat{T}}}$$

ومنها فإننا نجد:

$$Z'_{1-\frac{\alpha}{2}} = \frac{\left|\frac{\widehat{T} - T}{T}\right|}{\frac{\delta_{\widehat{T}}}{T}}$$

بتربيع الطرفين يصبح لدينا:

$$Z'^{2}_{1-\frac{\alpha}{2}} = \frac{p'^{2}}{\sum_{i=1}^{k} W_i^2 \frac{\delta_i^2}{n_i \mu^2} - \sum_{i=1}^{k} W_i \frac{\delta_i^2}{N\mu^2}}$$

$$p'^{2} = Z'^{2}_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i^2 . \frac{\delta_i^2}{n_i \mu^2} - Z'^{2}_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i . \frac{\delta_i^2}{N\mu^2}$$

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة على طبقات المجتمع $\left(n_i = \frac{n}{k}\right)$ فإن:

$$p'^{2} = k . Z'^{2}_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i^2 . \frac{\delta_i^2}{n\,\mu^2} - Z'^{2}_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i . \frac{\delta_i^2}{N\mu^2}$$

$$\boxed{n = \frac{k . \sum_{i=1}^{k} N_i^2 . \delta_i^2}{\frac{N^2 \mu^2 p'^2}{Z'^{2}_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i . \delta_i^2}}$$

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة على طبقات المجتمع $(n_i = n \,.\, W_i)$ فإن:

$$p'^{2} = Z'^{2}_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i . \frac{\delta_i^2}{n\mu^2} - Z'^{2}_{1-\frac{\alpha}{2}} . \sum_{i=1}^{k} W_i . \frac{\delta_i^2}{N\mu^2}$$

$$n = \frac{N.\sum_{i=1}^{k} N_i.\delta_i^2}{\frac{N^2\mu^2 p'^2}{Z'^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}$$

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة على طبقات المجتمع $(n_i = n.\frac{\frac{N_i\delta_i}{\sqrt{c_i}}}{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}})$ فإن:

$$N^2\mu^2 p'^2 = Z'^2_{1-\frac{\alpha}{2}}.\frac{\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}.\sum_{i=1}^{k} N_i\delta_i\sqrt{c_i}}{n} - Z'^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} N_i.\delta_i^2$$

$$n = \frac{(\sum_{i=1}^{k}\frac{N_i\delta_i}{\sqrt{c_i}}).(\sum_{i=1}^{k} N_i\delta_i\sqrt{c_i})}{\frac{N^2\mu^2 p'^2}{Z'^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_i.\delta_i^2}$$

ملاحظة:

- ✓ إذا كانت قيم تباين الطبقات δ_i^2 غير معلومة، نستخدم بدلها مقدراتها غير المتحيزة s_i^2.
- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Z \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

5. تقدير نسبة المفردات (P) التي تتصف بصفة معينة في المجتمع

كثيرا ما يهتم الباحث بتقدير نسبة السكان الذين يتصفون بصفة معينة باستخدام المعاينة العشوائية الطبقية. ولتقدير نسبة صفة معينة P_i في الطبقة i $(\widehat{P_i})$، نضع: $\widehat{P_i} = \frac{\sum_{j=1}^{n_i} x_{ij}}{n_i}$، حيث تأخذ x_{ij} القيمتين 0 أو1، الواحد إذا كانت المفردة j في الطبقة i تتصف بالصفة المعينة والصفر إذا لم تكن تتصف بها. وعليه من أجل تقدير نسبة صفة معينة P في المجتمع:

$$\widehat{P} = \frac{1}{N}\sum_{i=1}^{k} N_i\widehat{P_i} = \sum_{i=1}^{k} W_i\widehat{P_i}$$

حيث $\widehat{P}$ تعتبر المقدرة غير المتحيزة ل P أي أن: $E(\widehat{P}) = P$

و $V(\widehat{P}) = \delta_{\widehat{P}}^2 = \sum_{i=1}^{k} W_i^2\delta_{\widehat{P}_i}^2$، علما أن: $\delta_{\widehat{P}_i}^2 = \left(1 - \frac{n_i}{N_i}\right).\frac{P_iQ_i}{n_i}$.

1.5. مجال الثقة للنسبة P

إذا كان حجم العينة كبيرا وكانت قيمة $\delta_{\widehat{P}}^2$ غير معلومة فإن:

$\widehat{P} \sim N(P, \hat{\delta}_{\widehat{P}}^2)$، وبالتالي: $Y = \frac{\widehat{P}-P}{\hat{\delta}_{\widehat{P}}} \sim N(0,1)$.

مجال الثقة للنسبة P عند درجة معنوية α يمكن التوصل إليه كما يلي:

لدينا: $P(-y \leq Y \leq +y) = 1 - \alpha$

أي: $-y \leq Y \leq +y$

ومنه: $-y \leq \frac{\widehat{P}-P}{\hat{\delta}_{\widehat{P}}} \leq +y$

وعليه يكون مجال ثقة النسبة P عند مستوى ثقة (1-α) هو:

$$P \in \left[\widehat{P} - y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{P}} , \widehat{P} + y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{P}}\right]$$

ملاحظة:

- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

مثال 4-3:

لدراسة نسبة العمال الذين يملكون سيارات شخصية بمؤسسة ما والبالغ عددهم 7500 عامل، قررت الإدارة لهذا الغرض، وبالاعتماد على مبلغ الأجر الشهري لكل عامل، تقسيم مجموع العمال على 03 طبقات (أصحاب الدخل الضعيف- الطبقة 01، أصحاب الدخل المتوسط- الطبقة 02، أصحاب الدخل الكبير- الطبقة 03)، فكانت النتائج على النحو التالي:

الطبقة i	حجم الطبقة N_i	حجم العينة n_i	نسبة العمال الذين يملكون سيارات شخصية $\widehat{P}_i$
1	3.500	500	0,13
2	2.000	300	0,45
3	2.000	200	0,50
المجموع	7.500	1.000	

المطلوب:

- أوجد مجال الثقة لنسبة العمال الذين يملكون سيارات شخصية بالمؤسسة P وذلك عند درجة معنوية 5%.

الحل:

من خلال المعطيات نجد أن نسبة عمال المؤسسة الذين يملكون سيارات شخصية في العينة المختارة:

$$\widehat{P} = \sum_{i=1}^{3} W_i \widehat{P}_i = \frac{3500}{7500}.0{,}13 + \frac{2000}{7500}.0{,}45 + \frac{2000}{7500}.0{,}50 = 0{,}314$$

ويمكن حساب مقدرة تباين نسبة عمال المؤسسة الذين يملكون سيارات شخصية $\hat{\delta}^2_{\widehat{P}}$ حيث:

$\hat{\delta}^2_{\widehat{P}} = \sum_{i=1}^{3} W_i^2 \hat{\delta}^2_{\widehat{P}_i}$، علما أن: $\hat{\delta}^2_{\widehat{P}_i} = \left(1 - \frac{n_i}{N_i}\right).\frac{\widehat{P}_i \widehat{Q}_i}{n_i}$.

بالتعويض نجد:

$$\hat{\delta}^2_{\widehat{P}} = 0{,}000172,\ \hat{\delta}_{\widehat{P}} = \sqrt{0{,}000172} = 0{,}0131$$

من جدول التوزيع الطبيعي $N(0,1)$ نجد أن: $y_{1-\frac{\alpha}{2}} = 1{,}96$.

وعليه فإن مجال ثقة P وذلك عند مستوى ثقة 0.95 هو:

$$P \in \left[\widehat{P} - y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{P}}\ ,\ \widehat{P} + y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\widehat{P}}\right]$$

$$P \in [0{,}314 - (1{,}96).(0{,}0131)\ ,\ 0{,}314 + (1{,}96).(0{,}0131)]$$

$$P \in [0.288\ ,\ 0.339]$$

أي أن قيمة P لا يمكن أن تقل عن 0,288 ولا أن تزيد عن 0,339 وذلك عند مستوى ثقة 0.95.

2.5. تقدير حجم العينة اللازم لتقدير النسبة P بمستوى دقة f

بنفس الطريقة المعتمدة في تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ لدينا:

$$f = \left|\widehat{P} - P\right|$$

$$Y = \frac{\left|\widehat{P} - P\right|}{\delta_{\widehat{P}}}$$

بتربيع الطرفين يصبح لدينا:

$$Y^2_{1-\frac{\alpha}{2}} = \frac{f^2}{\sum_{i=1}^{k} W_i^2 \delta^2_{\widehat{P}_i}} = \frac{f^2}{\sum_{i=1}^{k} W_i^2 \left(1 - \frac{n_i}{N_i}\right).\frac{P_i Q_i}{n_i}}$$

$$Y^2_{1-\frac{\alpha}{2}} = \frac{f^2}{\sum_{i=1}^{k} W_i^2.\frac{P_i Q_i}{n_i} - \sum_{i=1}^{k} W_i.\frac{P_i Q_i}{N}}$$

$$f^2 = Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i^2.\frac{P_i Q_i}{n_i} - Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{P_i Q_i}{N}$$

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة على طبقات المجتمع $(n_i = \frac{n}{k})$ فإن:

$$f^2 = k.Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i^2.\frac{P_iQ_i}{n} - Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{P_iQ_i}{N}$$

$$\boxed{n = \frac{k.\sum_{i=1}^{k} N_i^2 P_iQ_i}{\frac{N^2f^2}{Y^2_{1-\frac{á}{2}}} + \sum_{i=1}^{k} N_iP_iQ_i}}$$

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة على طبقات المجتمع $(n_i = n\,.\,W_i)$ فإن:

$$f^2 = Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i^2.\frac{P_iQ_i}{n_i} - Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{P_iQ_i}{N}$$

$$f^2 = Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{P_iQ_i}{n} - Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} W_i.\frac{P_iQ_i}{N}$$

$$\boxed{n = \frac{N.\sum_{i=1}^{k} N_iP_iQ_i}{\frac{N^2f^2}{Y^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_iP_iQ_i}}$$

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة على طبقات المجتمع $(n_i = n\,.\,\frac{\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}}{\sum_{i=1}^{k}\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}})$ فإن:[20]

$$N^2f^2 = Y^2_{1-\frac{\alpha}{2}}.\frac{\sum_{i=1}^{k}\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}.\sum_{i=1}^{k} N_i\sqrt{P_iQ_i}.\sqrt{c_i}}{n} - Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k} N_i.P_iQ_i$$

$$\boxed{n = \frac{(\sum_{i=1}^{k}\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}).(\sum_{i=1}^{k} N_i\sqrt{P_iQ_i}.\sqrt{c_i})}{\frac{N^2f^2}{Y^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k} N_iP_iQ_i}}$$

ملاحظة:

✓ في حالة عدم توفر القيم الفعلية P_i وQ_i نستعمل مقدراتها $(\widehat{P_i}, \widehat{Q_i})$.

[20] Pascal ARDILLY, Op cit, Page 107.

✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y \sim t_{n-1}^{1-\frac{\alpha}{2}}$، بدل التوزيع الطبيعي $N(0,1)$.

3.5. تقدير حجم العينة اللازم لتقدير النسبة P بمستوى دقة نسبية h

بنفس الطريقة المعتمدة في تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ لدينا:

$$h = \left|\frac{\widehat{P} - P}{P}\right|$$

$$Y = \frac{\left|\frac{\widehat{P} - P}{P}\right|}{\frac{\delta_{\widehat{P}}}{P}}$$

بتربيع الطرفين نجد:

$$Y_{1-\frac{\alpha}{2}}^2 = \frac{h^2}{\frac{\sum_{i=1}^{k} W_i^2 . \frac{P_i Q_i}{n_i} - \sum_{i=1}^{k} W_i . \frac{P_i Q_i}{N}}{P^2}}$$

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة على طبقات المجتمع ($n_i = \frac{n}{k}$) فإن:

$$h^2 = k . \frac{Y_{1-\frac{\alpha}{2}}^2}{P^2} . \sum_{i=1}^{k} W_i^2 . \frac{P_i Q_i}{n} - \frac{Y_{1-\frac{\alpha}{2}}^2}{P^2} . \sum_{i=1}^{k} W_i . \frac{P_i Q_i}{N}$$

$$\boxed{n = \frac{k . \sum_{i=1}^{k} N_i^2 P_i Q_i}{\frac{N^2 P^2 h^2}{Y_{1-\frac{\alpha}{2}}^2} + \sum_{i=1}^{k} N_i P_i Q_i}}$$

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة على طبقات المجتمع ($n_i = n \, . \, W_i$) فإن:

$$h^2 = \frac{Y_{1-\frac{\alpha}{2}}^2}{P^2} . \sum_{i=1}^{k} W_i . \frac{P_i Q_i}{n} - \frac{Y_{1-\frac{\alpha}{2}}^2}{P^2} . \sum_{i=1}^{k} W_i . \frac{P_i Q_i}{N}$$

$$\boxed{n = \frac{N . \sum_{i=1}^{k} N_i P_i Q_i}{\frac{N^2 P^2 h^2}{Y_{1-\frac{\alpha}{2}}^2} + \sum_{i=1}^{k} N_i P_i Q_i}}$$

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة على طبقات المجتمع ($n_i = n \cdot \frac{\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}}{\sum_{i=1}^{k}\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}}$) فإن:

$$N^2P^2h^2 = Y^2_{1-\frac{\alpha}{2}}.\frac{\sum_{i=1}^{k}\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}.\sum_{i=1}^{k}N_i\sqrt{P_iQ_i}.\sqrt{c_i}}{n} - Y^2_{1-\frac{\alpha}{2}}.\sum_{i=1}^{k}N_iP_iQ_i$$

$$\boxed{n = \frac{(\sum_{i=1}^{k}\frac{N_i\sqrt{P_iQ_i}}{\sqrt{c_i}}).(\sum_{i=1}^{k}N_i\sqrt{P_iQ_i}.\sqrt{c_i})}{\frac{N^2P^2h^2}{Y^2_{1-\frac{\alpha}{2}}} + \sum_{i=1}^{k}N_iP_iQ_i}}$$

ملاحظة:

- ✓ في حالة عدم توفر القيم الفعلية P_i وQ_i نستعمل مقدراتها $(\widehat{P}_i, \widehat{Q}_i)$.
- ✓ إذا كان حجم العينة صغيرا ($n \leq 30$)، نستعمل توزيع ستيودنت (Student) بدرجة حرية (n-1) أي $Y \sim t^{1-\frac{\alpha}{2}}_{n-1}$، بدل التوزيع الطبيعي $N(0,1)$.

مثال 4-4:

بالعودة لبيانات المثال 4-3 وعلى اعتبار الدراسة بمثابة دراسة أولية، **المطلوب:**

1/ أوجد تقدير لحجم العينة العشوائية الطبقية اللازم سحبها لتقدير P وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق النسبي بين قيمة P الفعلية والمقدرة، $h = \left|\frac{\widehat{P}-P}{P}\right| = 0.2$ وذلك بدرجة معنوية 5%.

ملاحظة: تكاليف سحب المفردة الواحدة من كل طبقة هي كالآتي:

$c_1 = 1.50, c_2 = 1.75, c_3 = 2.$

الحل:

-حجم العينة العشوائية الطبقية اللازم سحبها لتقدير P وذلك باستخدام أساليب التخصيص المختلفة:

1/ باستخدام عبارة التوزيع المتساوي لحجم العينة نجد:

$$n = \frac{3.(3375475)}{\frac{(7500)^2.(0{,}314)^2.(0{,}2)^2}{(1{,}96)^2} + 1390{,}85} = 171{,}23$$

وعليه فإن حجم العينة اللازم سحبها هو $n = 172$ عامل من بينهم 80 من الطبقة 01، 46 من الطبقة 02 و46 من الطبقة 03 (بالنظر لوزن كل طبقة في المجتمع الكلي W_i).

2/ باستخدام عبارة التوزيع المتناسب لحجم العينة نجد:

$$n = \frac{7500.(1390,85)}{\frac{(7500)^2.(0,314)^2.(0,2)^2}{(1,96)^2} + 1390,85} = 176,39$$

وعليه فإن حجم العينة اللازم سحبها هو $n = 177$ عامل من بينهم 83 من الطبقة 01، 47 من الطبقة 02 و47 من الطبقة 03 (بالنظر لوزن كل طبقة في المجتمع الكلي W_i).

3/ باستخدام عبارة التوزيع الأمثل تحت قيد الميزانية لحجم العينة نجد:

$$n = \frac{(2420,31).(4172,06)}{\frac{(7500)^2.(0,314)^2.(0,2)^2}{(1,96)^2} + 1390,85} = 170,74$$

وعليه فإن حجم العينة اللازم سحبها هو $n = 171$ عامل من بينهم 80 من الطبقة 01، 46 من الطبقة 02 و45 من الطبقة 03 (بالنظر لوزن كل طبقة في المجتمع الكلي W_i).

6. مقارنة بين العينة العشوائية الطبقية والعينة العشوائية البسيطة

إن الهدف من المعاينة الطبقية هو تشكيل مجموعات جزئية متجانسة حيث يكون التباين بين الطبقات (variance inter-strates) أكبر ما يمكن، في المقابل يكون التباين بين مفردات كل طبقة على حدى (variance intra-strates) أصغر ما يمكن (تجانس كبير بين مفردات الطبقة الواحدة).

مما سبق وجدنا أن متوسط المجتمع يعطى بالعلاقة:

$$\bar{x} = \sum_{i=1}^{k} \frac{N_i}{N} \bar{x}_i = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} \frac{N_i}{n_i} x_{ij}$$

علما أن $f_i = \frac{n_i}{N_i}$ تمثل نسبة الاستقصاء أو إحتمال الظهور وتختلف قيمتها من طبقة إلى أخرى، وباستعمال أسلوب التخصيص المتناسب ($\frac{n}{N} = \frac{n_i}{N_i}$ أي $n_i = n.\frac{N_i}{N}$)، تصبح f_i ثابتة بالنسبة لكل الطبقات أي $f = \frac{n}{N} = \frac{n_i}{N_i} = f_i$ مما يعني أنه لكل مفردة نفس الوزن التمثيلي في العينة المختارة والذي يساوي $\frac{N}{n} = \frac{1}{f}$. وعليه فإن العلاقة السابقة تصبح: $\bar{x} = \frac{1}{n} \sum_{i=1}^{k} \sum_{j=1}^{n_i} x_{ij}$.

بتباين: $V(\bar{x}) = (1 - f) \frac{1}{n} \sum_{i=1}^{k} \frac{N_i}{N} \delta_i^2$.

هذه العلاقة الأخيرة توضح أنه كلما كان تجانس الطبقات كبيرا (قيم δ_i^2 صغيرة)، كلما كانت قيمة $V(\bar{x})$ صغيرة وبالتالي المعاينة الطبقية فعالة[21]. تجدر الإشارة إلى أن عبارة تباين متوسط المجتمع بالنسبة للمعاينة العشوائية البسيطة $\delta_{\bar{x}}^2$ يمكن كتابتها بالشكل:

$$\delta_{\bar{x}}^2 = V(\bar{x}) + (1-f)\frac{1}{n}\sum_{i=1}^{k}\frac{N_i}{N}(\bar{x}_i - \bar{x})^2$$

وبالتالي فإن: $V(\bar{x}) \leq \delta_{\bar{x}}^2$ (قيمة تباين متوسط المجتمع بالنسبة للمعاينة الطبقية لا تزيد عن قيمة تباين متوسط المجتمع بالنسبة للمعاينة البسيطة)، وبالتالي فالمعاينة الطبقية تعد أكثر فعالية من المعاينة البسيطة من حيث النتائج المتحصل عليها وبالتالي تعميمها.

[21] Yves TILLÉ, "Théorie des sondages : échantillonnage et estimation en population finie : cours et exercices avec solutions", Édition DUNOD, Paris 2011, Page 123.

سلسلة تمارين مقترحة حول المحور الرابع:

تمرين 01:

أثبت صحة علاقة التباين δ_x^2 الآتية:

$$\delta_x^2 = \frac{1}{N}\sum_{i=1}^{k}\sum_{j=1}^{N_i}(x_{ij} - \bar{x})^2 = \frac{1}{N}\sum_{i=1}^{k} N_i\delta_i^2 + \frac{1}{N}\sum_{i=1}^{k} N_i(\bar{x}_i - \bar{x})^2$$

تمرين 02:

لدراسة الاستهلاك الشهري للأسرة الواحدة من الكهرباء في إحدى البلديات المكونة من خمسة أحياء، وعلى اعتبار كل حي بمثابة طبقة تقرر الاعتماد على المعاينة العشوائية الطبقية لأجل هذه الدراسة، وفيما يلي جدول يلخص أهم البيانات التي تم جمعها لهذه الغاية:

الطبقة i	عدد الأسر N_i	متوسط إستهلاك العينة $\bar{x}_i$	تباين العينة s_i^2
1	400	10	6
2	250	20	12
3	200	20	15
4	100	30	20
5	50	35	30
المجموع	1.000		

المطلوب:

1/ أوجد مجال الثقة لمتوسط الاستهلاك الشهري الأسري في هذه البلدية μ عند درجة معنوية 5%.

2/ تقدير حجم العينة العشوائية الطبقية اللازم سحبها من هذه البلدية لتقدير μ وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق بين المتوسط الفعلي والمقدر، $d = |\bar{x} - \mu| = 1{,}5$ دينار وذلك بدرجة معنوية 5%.

3/ تقدير حجم العينة العشوائية الطبقية اللازم سحبها من هذه البلدية لتقدير μ وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق النسبي بين المتوسط الفعلي والمقدر، $p = \left|\frac{\bar{x}-\mu}{\mu}\right| = 0.08$ وذلك بدرجة معنوية 5%.

ملاحظة: تكاليف سحب المفردة الواحدة من كل طبقة هي كالآتي:

$c_1 = 1.20, c_2 = 1.25, c_3 = 1.35, c_4 = 1.40, c_5 = 1.50.$

تمرين 03:

بالاعتماد على بيانات التمرين رقم 02، المطلوب:

1/ أوجد مجال الثقة لإجمالي الاستهلاك الشهري للأسر في هذه البلدية T عند درجة معنوية 5%.

2/ تقدير حجم العينة العشوائية الطبقية اللازم سحبها من هذه البلدية لتقدير T وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق بين قيمة T الفعلية والمقدرة، $d' = 1000 = \left|\widehat{T} - T\right|$ دينار وذلك بدرجة معنوية 5%.

3/ تقدير حجم العينة العشوائية الطبقية اللازم سحبها من هذه البلدية لتقدير T وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق النسبي بين قيمة T الفعلية والمقدرة، $p' = \left|\frac{\widehat{T}-T}{T}\right| = 0.05$ وذلك بدرجة معنوية 5%.

تمرين 04:

لدراسة نسبة الزبائن الذين هم على استعداد للاكتتاب في المنتوج المالي الجديد لأحد البنوك، تم تقسيم الزبائن حسب الأرصدة المالية الموجودة في حساباتهم إلى ثلاثة طبقات على النحو التالي (الأرصدة الصغيرة، المتوسطة والكبيرة)، سحبنا من كل طبقة عينة عشوائية من الزبائن وكانت البيانات المتحصل عليها كالآتي:

الطبقة i	حجم الطبقة N_i	حجم العينة n_i	النسبة في العينة $\widehat{P}_i$
1	2.500	250	0.08
2	1.500	150	0.10
3	1.000	100	0.12
المجموع	5.000	500	

المطلوب:

1/ أوجد تقدير لنسبة الزبائن المستعدين للاكتتاب في المنتوج المالي الجديد للبنك $\widehat{P}$.

2/ أوجد تقدير لتباين نسبة الزبائن المستعدين للاكتتاب في المنتوج المالي الجديد للبنك $\delta^2_{\widehat{P}}$.

3/ أوجد مجال الثقة لنسبة الزبائن المستعدين للاكتتاب في المنتوج المالي الجديد للبنك P وذلك عند درجة معنوية 5%.

4/ أوجد تقدير لحجم العينة العشوائية الطبقية اللازم سحبها لتقديرP وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق بين قيمة P الفعلية والمقدرة، $f = 0.02 = \left|\widehat{P} - P\right|$ وذلك بدرجة معنوية 5%.

5/ أوجد تقدير لحجم العينة العشوائية الطبقية اللازم سحبها لتقديرP وذلك باستخدام أساليب التخصيص المختلفة، بحيث لا يتعدى الفرق المطلق النسبي بين قيمة P الفعلية والمقدرة، $h = 0.2 = \left|\frac{\widehat{P}-P}{P}\right|$ وذلك بدرجة معنوية 5%.

ملاحظة: تكاليف سحب المفردة الواحدة من كل طبقة هي كالآتي:

$c_1 = 1.50, c_2 = 2, c_3 = 2.50.$

المحور الخامس: العينة العشوائية المنتظمة

إن استخدام المعاينة العشوائية المنتظمة يقوم في الأساس على مدى إمكانية توفر مجتمع دراسة قابل للترتيب إما تصاعديا وإما تنازليا، ونميز عادة بين أسلوبين لسحب مفردات العينة وهما الأسلوب الخطي المنتظم والأسلوب الدائري المنتظم، فالأسلوب الأول وهو الأكثر استعمالا والذي يتطلب ترقيم مشاهدات المجتمع ثم اختيار مشاهدة عشوائيا من أول $(k = \frac{N}{n})$ مشاهدة ولتكن المفردة h، وإذا أردنا اختيار المشاهدة الثانية فإننا ننتقل إلى المفردة $(h + k)$ وهكذا بالنسبة لبقية مشاهدات العينة، حتى نحصل على كل المفردات (h+ik) حيث $i = \overline{0, (n-1)}$. ويستخدم هذا الأسلوب عادة عندما تكون N من مضاعفات n حتى لا يتم تقريب قيمة k إلى أقرب عدد صحيح، والذي قد ينجم عنه إختيار عينات لا يساوي حجمها n مما ينجر عنه تقديرات متحيزة، وهذا ما يضطرنا للجوء إلى الأسلوب الثاني.

فأسلوب السحب الدائري المنتظم يسمح لنا باختيار المفردة الأولى من بين كل مفردات المجتمع والتي عددها N ولتكن المفردة h، وإذا أردنا اختيار المشاهدة الثانية فإننا ننتقل إلى المفردة $(h + k)$، حيث يمثل k أقرب عدد صحيح لقسمة $\frac{N}{n}$، وهكذا بالنسبة لبقية مشاهدات العينة، حتى نحصل على كل المفردات:

- ✓ (h+ik) إذا كانت $h + ik \leq N$ ، حيث $i = \overline{0, (n-1)}$.
- ✓ (h+ik-N) إذا كانت $h + ik > N$ ، حيث $i = \overline{0, (n-1)}$.[22]

1. مجال الثقة بالنسبة لمتوسط المجتمع μ والمجموع الكلي T

يعطى الوسط الحسابي للعينة العشوائية المنتظمة $\bar{x}$ بالعلاقة:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

وتباين: $\delta_{\bar{x}}^2 = \frac{\delta^2}{n}(1 + (n-1)\rho)$ عندما تكون المعاينة من مجتمع غير محدود، حيث تمثل ρ معامل الارتباط الذاتي بين كل مفردتين من مفردات العينة. في المقابل، عندما تكون المعاينة من مجتمع محدود حجمه N وبسحب من دون إعادة فإن:[23]

$$\delta_{\bar{x}}^2 = (1 - \frac{n}{N})\frac{\delta^2}{n}(1 + (n-1)\rho)$$

من خلال هذه العبارة الأخيرة يتضح أن قيمة ρ لا يمكن أن تكون كبيرة إشارة السالبة لأن قيمة $\delta_{\bar{x}}^2$ ستكون سالبة وهذا غير مقبول.

---

[22] فتحي أحمد عاروري، مرجع سبق ذكره، ص211.

[23] عدنان عباس حميدان، "تقنيات المعاينة الاحصائية في الدراسات السوقية والتسويقية"، دار الإعصار العلمي للنشر والتوزيع، عمان 2015، ص376.

أين يمكن إستعمال s^2 مقدرة غير متحيزة لتباين المجتمع δ^2 حيث:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

أو $s^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ إذا كانت n كبيرة (n>30).

في المقابل، تعطى ρ بالعبارة:

$$\rho = \frac{2}{(N-1)(n-1)\delta^2}\sum_{i=1}^{k}\sum_{j\neq j'=1}^{n}(x_{ij} - \bar{x})(x_{ij'} - \bar{x})$$

حيث: x_{ij} تمثل المفردة j $(j = \overline{1,n})$ في العينة i $(i = \overline{1,k})$.

: $x_{ij'}$ تمثل المفردة j' $(\neq j)$ في العينة i $(i = \overline{1,k})$.

: k تمثل عدد العينات الممكن تشكيلها بأسلوب المعاينة العشوائية المنتظمة.

مثال 5-1:

لتكن علامات 09 طلبة في مقياس معين كما يلي:

17,16,14,15,17,11,13,12,10

فيما يلي جدول يوضح مختلف العينات ذوات الحجم $n = 3$ الممكن سحبها باستعمال أسلوب المعاينة العشوائية المنتظمة:

رقم المفردة في العينة j $(j = \overline{1,3})$	العينة الأولى x_{1j}	العينة الثانية x_{2j}	العينة الثالثة x_{3j}
1	17	16	14
2	15	17	11
3	13	12	10
المجموع	45	45	35

المطلوب:

1/ أحسب معامل الارتباط الذاتي ρ.

2/ أحسب تباين الوسط الحسابي $\delta_{\bar{x}}^2$.

الحل:

1/ يحسب معامل الارتباط الذاتي ρ كما يلي:

$\bar{x} = \frac{1}{N}\sum_{i=1}^{3}\sum_{j=1}^{3}x_{ij} = \frac{1}{9}(125) = 13{,}88.$

$$\sum_{i=1}^{3}\sum_{j\neq j'=1}^{3}\left(x_{ij}-\bar{x}\right)\left(x_{ij'}-\bar{x}\right)=$$

$$=(17-13{,}88)(15-13{,}88)+(17-13{,}88)(13-13{,}88)$$
$$+(15-13{,}88)(13-13{,}88)+(16-13{,}88)(17-13{,}88)$$
$$+(16-13{,}88)(12-13{,}88)+(17-13{,}88)(12-13{,}88)$$
$$+(14-13{,}88)(11-13{,}88)+(14-13{,}88)(10-13{,}88)$$
$$+(11-13{,}88)(10-13{,}88)$$

$\sum_{i=1}^{3}\sum_{j\neq j'=1}^{3}\left(x_{ij}-\bar{x}\right)\left(x_{ij'}-\bar{x}\right)=6{,}89.$

$\delta^2=\frac{1}{N-1}\sum_{i=1}^{k}\sum_{j\neq j'=1}^{n}\left(x_{ij}-\bar{x}\right)^2=\frac{1}{9-1}(52{,}89)=6{,}61.$

$\rho=\frac{2}{(9-1).(3-1).(6{,}61)}6{,}89=0{,}13.$

2/ تباين الوسط الحسابي:

$\delta_{\bar{x}}^2=\left(1-\frac{3}{9}\right)\frac{6{,}61}{3}\left(1+(3-1).(0{,}13)\right)=1{,}08.$

1.1. مجال الثقة بالنسبة لمتوسط المجتمع μ

كما رأينا سابقا في حالتي المعاينة البسيطة والطبقية، فإن مجال الثقة بالنسبة لمتوسط المجتمع μ غير المعلوم بمستوى معنوية α يمكن حسابه كما يلي:

مع العلم أن:

$$P(-z\leq Z\leq +z)=1-\alpha$$

أي:

$$-z\leq Z\leq +z$$

$$-z\leq\frac{\bar{x}-\mu}{\delta_{\bar{x}}}\leq +z$$

وعليه يكون مجال الثقة بالنسبة ل μ بمستوى ثقة $(1-\alpha)$ هو:

$$\boxed{\mu\in\left[\bar{x}-z_{1-\frac{\alpha}{2}}.\delta_{\bar{x}}\,,\bar{x}+z_{1-\frac{\alpha}{2}}.\delta_{\bar{x}}\right]}$$

كما يمكن إحتساب مجال الثقة للمتوسط μ أيضا باستخدام طريقة حد الخطأ في التقدير، حيث:

$$B=2.\delta_{\bar{x}}$$

أي أن متوسط المجتمع μ، وذلك عند مستوى ثقة $(1-\alpha)$ يقع في المدى (مجال الثقة) الآتي:

$$\boxed{\mu=\bar{x}\mp B}$$

2.1. مجال الثقة بالنسبة للمجموع الكلي T

علما أن تقدير المجموع الكلي T هو $\widehat{T}=N\bar{x}$ ، بمتوسط: $E(\widehat{T})=N\mu=T$،

وتباين: $V(\widehat{T}) = N^2(1 - \frac{n}{N})\frac{\delta^2}{n}(1 + (n-1)\rho)$.

فإن مجال الثقة بالنسبة للمجموع الكلي للمجتمع T غير المعلوم بمستوى معنوية α يمكن حسابه كما يلي:

$$P(-z' \leq Z' \leq +z') = 1 - \alpha$$

أي:

$$-z' \leq Z' \leq +z'$$

$$-z' \leq \frac{\widehat{T} - T}{\delta_{\widehat{T}}} \leq +z'$$

$$\widehat{T} - z'.\delta_{\widehat{T}} \leq T \leq \widehat{T} + z'.\delta_{\widehat{T}}$$

$$T \in \left[\widehat{T} - z'.\delta_{\widehat{T}}\,,\widehat{T} + z'.\delta_{\widehat{T}}\right]$$ أو

وعليه يكون مجال الثقة بالنسبة ل T بمستوى ثقة (1-α) هو:

$$\boxed{T \in \left[\widehat{T} - z'_{1-\frac{\alpha}{2}}.\delta_{\widehat{T}}\,,\widehat{T} + z'_{1-\frac{\alpha}{2}}.\delta_{\widehat{T}}\right]}$$

كما يمكن إحتساب مجال الثقة للمجموع الكلي T أيضا باستخدام طريقة حد الخطأ في تقدير المجموع T ($\widehat{T}$)، حيث:

$$B' = 2.\delta_{\widehat{T}}$$

أي أن المجموع الكلي T، وذلك عند مستوى ثقة $(1-\alpha)$ يقع في المدى (مجال الثقة) الآتي:

$$\boxed{T = \widehat{T} \mp B'}$$

مثال 5-2:

لأجل دراسة متوسط عدد عمال الشركات المتواجدة على مستوى إحدى المناطق الصناعية، تم سحب عينة عشوائية منتظمة حجمها n=5 شركة من مجموع الشركات التي يقدر عددها N=15. إذا كانت الشركات مرتبة حسب حجم العمالة وتم اختيار الشركة رقم 02 كعنصر بداية للعينة، وأمكن اختيار العينة العشوائية المنتظمة المكونة من الشركات الآتية أرقامهم:

02,05,08,11,14

وكان عدد العمال في كل شركة كما يلي:

300,285,250,230,190

المطلوب:

1/ أوجد مجال الثقة لمتوسط عدد عمال الشركات المتواجدة على مستوى المنطقة الصناعية μ وذلك عند درجة معنوية 5%.

2/ أوجد مجال الثقة لإجمالي عدد عمال الشركات المتواجدة على مستوى المنطقة الصناعية T وذلك عند درجة معنوية 5%.

الحل:

1/ مجال الثقة ل μ:

في حالة عدم توفر قيمة معامل الارتباط الذاتي ρ وهو ما يعتبر واقع الحال في كثير من الأحيان، فإننا من الناحية العملية نستخدم في المعاينة العشوائية المنتظمة نفس القواعد المستخدمة في المعاينة العشوائية البسيطة:[24]

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = \frac{1}{5}(300 + 285 + 250 + 230 + 190) = 251,$$
$$s^2 = \frac{1}{n-1}\sum_{i=1}^{5}(x_i - \bar{x})^2 = \frac{1}{4}(7720) = 1930,$$
$$\hat{\delta}_{\bar{x}}^2 = \left(1 - \frac{n}{N}\right)\frac{s^2}{n} = \left(1 - \frac{5}{15}\right)\frac{1930}{5} = 257{,}33.$$
$$\hat{\delta}_{\bar{x}} = \sqrt{257{,}33} = 16{,}04.$$

من جدول التوزيع الطبيعي $N(0,1)$ نجد أن: $z_{1-\frac{\alpha}{2}} = 1{,}96$.

وعليه فإن مجال ثقة μ وذلك عند مستوى ثقة 0.95 هو:

$$\mu \in [251 - (1{,}96).(16{,}04), 251 + (1{,}96).(16{,}04)]$$
$$\mu \in [219.56\ , 282.43]$$

أي أن قيمة μ لا يمكن أن تقل عن 220 عامل ولا أن تزيد عن 283 عامل وذلك عند مستوى ثقة 0.95.

2/ مجال الثقة ل T:

$$\hat{T} = N\bar{x} = 15.(251) = 3765,$$
$$\hat{\delta}_{\hat{T}} = N.\hat{\delta}_{\bar{x}} = 15.(16{,}04) = 240{,}6.$$

وعليه فإن مجال ثقة T وذلك عند مستوى ثقة 0.95 هو:

$$T \in [3765 - (1{,}96).(240{,}6), 3765 + (1{,}96).(240{,}6)]$$
$$T \in [3293.42\ , 4236.57]$$

أي أن قيمة T لا يمكن أن تقل عن 3294 عامل ولا أن تزيد عن 4237 عامل وذلك عند مستوى ثقة 0.95.

2. تقدير حجم العينة اللازم لتقدير μ وT

1.2. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ

لتقدير حجم العينة n اللازم سحبها بالنسبة للعينة العشوائية الطبقية وذلك باستخدام طريقة حد الخطأ في تقدير μ حيث:

[24] فتحي أحمد عاروري، مرجع سبق ذكره، ص216.

$.B = 2.\delta_{\bar{x}}, D = \frac{B^2}{4}$

وبنفس الطريقة المستعملة في العينة العشوائية البسيطة نجد:

$$n = \frac{N\delta^2}{ND + \delta^2}$$

2.2. تقدير حجم العينة اللازم لتقدير المجموع الكلي T

لتقدير حجم العينة n اللازم سحبها بالنسبة للعينة العشوائية الطبقية وذلك باستخدام طريقة حد الخطأ في تقدير المجموع T حيث:

$.B' = 2.\delta_{\hat{T}} = 2.N.\delta_{\bar{x}}, D' = \frac{B'^2}{4N^2}$

وبنفس الطريقة المستعملة سابقا في تقدير حجم العينة اللازم سحبها لتقدير μ نجد:

$$n = \frac{N\delta^2}{ND' + \delta^2}$$

مثال 5-3:

بالعودة لبيانات المثال 5-2، **المطلوب:**

-تقدير حجم العينة العشوائية المنتظمة اللازم سحبها لتقدير متوسط عدد عمال الشركات المتواجدة على مستوى المنطقة الصناعية، بحيث لا يتعدى حد الخطأ في التقدير ال $B = 34$.

الحل:

- حجم العينة العشوائية المنتظمة اللازم سحبها:

$$D = \frac{B^2}{4} = \frac{(34)^2}{4} = 289,$$

$$n = \frac{NS^2}{ND+S^2} = \frac{15.(1930)}{15.(289)+1930} = 4{,}62.$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 5$.

3. تقدير نسبة المفردات (P) التي تتصف بصفة معينة في المجتمع

لتقدير نسبة المفردات التي تتصف بصفة معينة باستخدام المعاينة العشوائية الطبقية $P\ (\widehat{P})$، نضع: $\widehat{P} = \frac{\sum_{i=1}^{n} x_i}{n} = \frac{r}{n}$، حيث تأخذ x_i القيمتين 0 أو1، الواحد إذا كانت المفردة تتصف بالصفة المعينة والصفر إذا لم تكن تتصف بها.

حيث $\widehat{P}$ تعتبر المقدرة غير المتحيزة ل P أي أن: $E(\widehat{P}) = P$،

و$V(\widehat{P}) = \delta^2_{\widehat{P}} = \frac{\widehat{P}\widehat{Q}}{n-1}(1 - \frac{n}{N})$.

1.3. مجال الثقة للنسبة P

إن مجال الثقة للنسبة P غير المعلومة بمستوى معنوية α يمكن حسابه كما يلي:

لدينا: $$P(-y \leq Y \leq +y) = 1 - \alpha$$

أي: $$-y \leq Y \leq +y$$

ومنه: $$-y \leq \frac{\hat{P}-P}{\hat{\delta}_{\hat{P}}} \leq +y$$

وعليه يكون مجال ثقة النسبة P عند مستوى ثقة (1-α) هو:

$$\boxed{P \in \left[\hat{P} - y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\hat{P}}\ , \hat{P} + y_{1-\frac{\alpha}{2}}.\hat{\delta}_{\hat{P}}\right]}$$

كما يمكن إحتساب مجال الثقة للنسبة P أيضا باستخدام طريقة حد الخطأ في تقدير P $(\hat{P})$، حيث:

$$B^* = 2.\delta_{\hat{P}}$$

أي أن النسبة P، وذلك عند مستوى ثقة $(1-\alpha)$ تقع في المدى (مجال الثقة) الآتي:

$$\boxed{P = \hat{P} \mp B^*}$$

2.3. تقدير حجم العينة اللازم لتقدير النسبة P

لتقدير حجم العينة n اللازم سحبها بالنسبة للعينة العشوائية الطبقية وذلك باستخدام طريقة حد الخطأ في تقدير النسبة P حيث:

$$.B^* = 2.\delta_{\hat{P}},\ D^* = \frac{B^{*2}}{4}$$

وبنفس الطريقة المستعملة سابقا في تقدير حجم العينة اللازم سحبها لتقدير μ نجد:

$$\boxed{n = \frac{NPQ}{ND^* + PQ}}$$

مثال 5-4:

لدراسة نسبة العمال الذين يعانون من مشكل السكن بإحدى المؤسسات الانتاجية التي تضم N=150 عامل، سحبنا عينة عشوائية منتظمة حجمها n=30 عامل، ومن خلالها تبين لنا أن r=10 عمال منهم يعانون من هذه المشكل، المطلوب:

1/ أوجد تقدير لنسبة العمال الذين يعانون من مشكل السكن بهذه المؤسسة P وذلك بمستوى ثقة 0.95.

2/ ماهو حجم العينة العشوائية المنتظمة اللازم سحبها لتقدير نسبة العمال الذين يعانون من مشكل السكن بهذه المؤسسة بحيث لا يتعدى حد الخطأ في تقدير النسبة P ال $B = 0,15$ ؟

الحل:

1/ مجال ثقة P:

$$\widehat{P} = \frac{r}{n} = \frac{10}{30} = 0{,}33.\ \widehat{Q} = 1 - \widehat{P} = 1 - 0{,}33 = 0{,}67.$$

$$\hat{\delta}^2_{\widehat{P}} = \frac{\widehat{P}\widehat{Q}}{n-1}\left(1 - \frac{n}{N}\right) = \frac{(0{,}33).(0{,}67)}{30-1}\left(1 - \frac{30}{150}\right) = 0{,}0061.$$

$$\hat{\delta}_{\widehat{P}} = \sqrt{0{,}0061} = 0{,}078.$$

من جدول التوزيع الطبيعي $N(0,1)$ نجد أن: $z_{1-\frac{\alpha}{2}} = 1{,}96$.

وعليه فإن مجال ثقة P وذلك عند مستوى ثقة 0.95 هو:

$$P \in [0{,}33 - (1{,}96).(0{,}078), 0{,}33 + (1{,}96).(0{,}078)]$$

$$P \in [0{,}18\ , 0{,}48]$$

أي أن قيمة P لا يمكن أن تقل عن 0,18 ولا أن تزيد عن 0,48 وذلك عند مستوى ثقة 0.95.

2/ حجم العينة اللازم سحبها:

$$D^* = \frac{B^{*2}}{4} = \frac{(0{,}15)^2}{4} = 0{,}005625.$$

$$n = \frac{NPQ}{ND^* + PQ} = \frac{150.(0{,}33).(0{,}67)}{150.(0{,}005625) + (0{,}33).(0{,}67)} = 31{,}15.$$

وعليه يكون حجم العينة اللازم سحبها هو $n = 32$. وبالتالي يجب على المؤسسة إضافة مفردتين أخريين للعينة المسحوبة سابقا.

4. مقارنة بين المعاينة العشوائية المنتظمة وكل من المعاينة العشوائية البسيطة والطبقية

بمقارنة العينتين العشوائيتين المنتظمة والطبقية فإنه يمكن اعتبار المعاينة المنتظمة بأسلوب السحب الخطي حالة خاصة من المعاينة العشوائية الطبقية. في المقابل، وللمقارنة بين المعاينة المنتظمة والبسيطة نلجأ للمقارنة بين قيمتي تباين الوسط الحسابي في الحالتين. وجدنا فيما سبق أن تباين الوسط الحسابي للعينة العشوائية البسيطة $(\bar{x})_{si}$ هو:

$$V(\bar{x})_{si} = (1 - \frac{n}{N})\frac{\delta^2}{n}$$

وأن تباين الوسط الحسابي للعينة العشوائية المنتظمة $(\bar{x})_{sy}$ هو:

$$V(\bar{x})_{sy} = (1 - \frac{n}{N})\frac{\delta^2}{n}(1 + (n-1)\rho)$$

وجه الاختلاف بين العلاقتين يتعلق بقيمة معامل الارتباط الذاتي ρ:

- ✓ إذا كانت قيمة ρ قريبة من الواحد، أي كان هناك تجانس وارتباط كبير بين مفردات العينة المنتظمة، فإن قيمة $V(\bar{x})_{sy} > V(\bar{x})_{si}$، وعليه يفضل استخدام المعاينة البسيطة.

- ✓ إذا كانت إشارة ρ سالبة وقيمة $(n-1)\rho$ أصغر تماما من الواحد، فإن قيمة $V(\bar{x})_{sy} < V(\bar{x})_{si}$، وعليه يفضل استخدام المعاينة المنتظمة.
- ✓ إذا كانت قيمة ρ معدومة، فإن قيمة $V(\bar{x})_{sy} = V(\bar{x})_{si}$، وعليه يمكن استخدام كلا الطريقتين لسحب العينة اللازمة للدراسة.

سلسلة تمارين مقترحة حول المحور الخامس:

تمرين 01:

قارن بين العينة العشوائية المنتظمة والعينة العشوائية البسيطة. ومتى نفضل طريقة على أخرى؟

تمرين 02:

إذا أرادت مؤسسة خاصة دراسة وجهة نظر عمالها إزاء النظام التعويضي المطبق على مستواها، وقررت لهذه الغاية سحب عينة عشوائية منتظمة حجمها n=20 عامل من مجموع N=200 عامل في المؤسسة مرتبين في قائمة خاصة على مستوى مصلحة المستخدمين.

المطلوب:

1/ ماهو أسلوب السحب المنتظم المطبق في هذه الحالة ؟

2/ ماهو عدد العينات الممكنة السحب في هذه الحالة ؟

3/ إذا علمت أنه قد تم إختيار العامل رقم 05 كعنصر بداية للعينة، أذكر العينة المختارة.

4/ إذا علمت أنه قد تم إختيار العامل رقم 07 كعنصر بداية للعينة، أذكر العينة المختارة.

تمرين 03:

إذا أردنا سحب عينة عشوائية منتظمة حجمها n=4 من مجتمع حجمه N=18، المطلوب:

1/ أوجد جميع العينات العشوائية المنتظمة الممكن سحبها باستخدام أسلوب السحب الخطي المنتظم (ضعها في جدول).

2/ ماهو عيب أسلوب السحب المذكور في هذه الحالة ؟ وماهو البديل ؟

3/ إذا علمت أنه قد تم إختيار المفردة رقم 08 كعنصر بداية للعينة، وذلك باستخدام أسلوب السحب الدائري المنتظم، أذكر العينة المختارة.

4/ إذا علمت أنه قد تم إختيار المفردة رقم 09 كعنصر بداية للعينة، وذلك باستخدام أسلوب السحب الدائري المنتظم، أذكر العينة المختارة.

تمرين 04:

إذا كانت علامات طلبة الماستر تخصص تسويق مصرفي، وعددهم N=15، المسجلين بكلية الاقتصاد في مادة تقنيات المعاينة كما يلي:

17,18,16,11,13,12,10,14,15,17,12,16,11.5,14.5,18

وأردنا أن نسحب منهم عينة عشوائية منتظمة حجمها n=5، المطلوب:

1/ ماهو أسلوب السحب المنتظم المستخدم في هذه الحالة ؟

2/ أوجد جميع العينات العشوائية المنتظمة الممكنة السحب باستخدام هذا الأسلوب (ومن خلال الترتيب السابق).

3/ أحسب معامل الارتباط الذاتي ρ.

تمرين 05:

لأجل دراسة حجم الساعات الإضافية التي يقوم بها أساتذة إحدى الكليات أسبوعيا، قرر الباحث سحب عينة عشوائية منتظمة حجمها n=20 أستاذ من مجموع أساتذة الكلية والذين يقدر عددهم N=200. إذا كانت قائمة الأساتذة مرتبة حسب الترتيب الأبجدي وتم اختيار الأستاذ رقم 07 كعنصر بداية للعينة، وأمكن اختيار العينة العشوائية المنتظمة المكونة من الأساتذة الآتية أرقامهم:

07,17,27,37,47,57,67,77,87,97,107,117,127,137,147,157,167,177,187,197

وكانت الساعات الاضافية الموافقة لكل أستاذ في العينة كما يلي:

4.5,9,7.5,12,15,9,0,6,18,12,7.5,4.5,6,10.5,15,10,5,3,0,3,18

المطلوب:

1/ أوجد مجال الثقة لمتوسط الساعات الإضافية التي يقوم بها أساتذة الكلية أسبوعيا μ وذلك عند درجة معنوية 5%.

2/ إختبر الفرضية القائلة بأن متوسط الساعات الإضافية التي يقوم بها أساتذة الكلية أسبوعيا لا يقل عن 12 ساعة وذلك عند درجة معنوية 5%.

3/ تقدير حجم العينة العشوائية المنتظمة اللازم سحبها لتقدير متوسط الساعات الإضافية التي يقوم بها أساتذة الكلية أسبوعيا، بحيث لا يتعدى الفرق المطلق بين قيمة μ الفعلية والمقدرة، $d = |\bar{x} - \mu| = 2$ ساعتين وذلك بدرجة معنوية 5%.

تمرين 06:

بالعودة لبيانات التمرين رقم 04 وعلى اعتبار هذه الدراسة بمثابة دراسة أولية:

1/ أوجد مجال الثقة لإجمالي الساعات الإضافية التي يقوم بها أساتذة الكلية أسبوعيا T وذلك عند درجة معنوية 5%.

2/ إختبر الفرضية القائلة بأن إجمالي الساعات الإضافية التي يقوم بها أساتذة الكلية أسبوعيا لا يزيد عن 5.000 ساعة وذلك عند درجة معنوية 5%.

3/ تقدير حجم العينة العشوائية المنتظمة اللازم سحبها لتقدير إجمالي الساعات الإضافية التي يقوم بها أساتذة الكلية أسبوعيا، بحيث لا يتعدى الفرق المطلق بين قيمة T الفعلية والمقدرة، $d' = \left|\widehat{T} - T\right| = 400$ ساعة وذلك بدرجة معنوية 5%.

تمرين 07:

لدراسة نسبة الطلبة الذين يعانون من مشاكل سمعية في قسم الرياضة، سحبنا عينة عشوائية منتظمة حجمها n=50 طالب، ومن خلالها تبين لنا أن r=5 طالب منهم يعانون من هذه المشاكل، فإذا علمنا بأن عدد طلبة قسم الرياضة هو N=250 طالب، المطلوب:

1/ أوجد تقدير لنسبة الطلبة الذين يعانون من مشاكل سمعية في قسم الرياضة وذلك بمستوى ثقة 0.95.

2/ إختبر الفرضية القائلة بأن نسبة الطلبة الذين يعانون من مشاكل سمعية في قسم الرياضة لا يقل عن 0.2 وذلك بمستوى ثقة 0.95.

3/ ماهو حجم العينة العشوائية المنتظمة اللازم سحبها لتقدير نسبة الطلبة الذين يعانون من مشاكل سمعية في قسم الرياضة بحيث لا يتعدى الفرق المطلق بين النسبة الحقيقة P والنسبة المقدرة $\widehat{P}$ ال $f = \left|\widehat{P} - P\right| = 0.08$، وذلك بمستوى ثقة 0.95 ؟

المحور السادس: العينة العشوائية العنقودية من مرحلة واحدة

يتم اختيار العينة وفق هذا الأسلوب على عدة مراحل، ويستخدم هذا الأسلوب في حالة ما إذا كان حجم المجتمع كبيرا جدا ولوجود قيود متعلقة بالإمكانات المتاحة للباحث من حيث الجهد، الوقت والمال. فيتم تقسيم المجتمع إلى وحدات تسمى بالوحدات الأولية أو العناقيد (grappes)، ثم يتم اختيار عينة عشوائية بسيطة من هذه الوحدات الأولية، ونقوم كذلك بتقسيم الوحدات المختارة إلى وحدات ثانوية ثم اختيار عينة عشوائية بسيطة منها وهكذا نستمر بالتقسيم والاختيار إلى أن نصل إلى العينة الأخيرة التي تحوي على مفردات تستنبط منها المعلومات المطلوبة.

تجدر الإشارة إلى أن وجه الاختلاف بين العينة العشوائية العنقودية والعينة العشوائية الطبقية يكمن في أن المفردات المكونة لكل طبقة بالنسبة للمعاينة الطبقية تكون متجانسة فيما بينها بالرغم من الاختلاف الموجود بين هذه الطبقات (طبقات غير متجانسة). في المقابل، فإن العناقيد بالنسبة للعينة العشوائية العنقودية تكون متجانسة فيما بينها رغم أن كل عنقود يحوي على مفردات قد تكون غير متجانسة (على سبيل المثال الوحدة الاحصائية "العائلة" وهي المكونة من أفراد مختلفين: الأب، الأم، الإبن، البنت...).

إذا كان المجتمع مقسما لوحدات أولية (عناقيد) ولنفترض أن عددها G، ونقوم باختيار عينة عشوائية بسيطة من هذه العناقيد وليكن عددها g، فإننا نكون أمام عينة عشوائية عنقودية من مرحلة واحدة فقط.

في هذه الحالة، مجتمع الدراسة هو N حيث: $N = \sum_{i=1}^{G} N_i$ مع $(i = 1,2, \dots, G)$.

والعينة المسحوبة هي كل مفردات العناقيد المختارة وليكن حجمها n حيث:

$$n = \sum_{i=1}^{g} N_i$$

أين: $(i = 1,2, \dots, g)$

N_i تمثل حجم العنقود المختار i.

1. مجال الثقة بالنسبة لمتوسط المجتمع μ والمجموع الكلي T

على إعتبار X_{ij} تمثل قيمة المفردة j في العنقود i، فإن الوسط الحسابي للعنقود μ_i يعطى بالعلاقة:

$$\mu_i = \frac{1}{N_i} \sum_{j=1}^{N_i} X_{ij}$$

أما تقدير الوسط الحسابي للمجتمع μ $(\overline{X}_g)$ فيعطى بالعلاقة:

$$\overline{X}_g = \frac{T}{n} = \frac{\sum_{i=1}^{g} T_i}{\sum_{i=1}^{g} N_i} = \frac{\sum_{i=1}^{g} \sum_{j=1}^{N_i} X_{ij}}{\sum_{i=1}^{g} N_i}$$

حيث: $T_i = \sum_{j=1}^{N_i} X_{ij}$ وتمثل المجموع الكلي للعنقود i.

كما تعطى عبارة تباين الوسط الحسابي المقدر $(\overline{X}_g)$ بالعلاقة:

$$\delta^2_{\overline{X}_g} = \left(\frac{G-g}{g.G.\overline{N}^2}\right)\frac{\sum_{i=1}^{g}(T_i - N_i\overline{X}_g)^2}{g-1}$$

حيث $\overline{N} = \frac{N}{G}$ وتمثل متوسط عدد المفردات في العنقود الواحد من مجموع العناقيد المكونة للمجتمع الكلي. نشير هنا إلى أنه في حالة عدم توفر قيمة $\overline{N}$، فإنه يمكن إستخدام مقدرتها $\bar{n} = \frac{n}{g}$ في العلاقة السابقة.[25]

1.1. مجال الثقة لمتوسط المجتمع μ

كما رأينا سابقا بالنسبة لطرق المعاينة المختلفة، يمكن إيجاد مجال الثقة للمتوسط μ بالنسبة للعينة العنقودية من مرحلة واحدة أيضا باستخدام طريقة حد الخطأ في التقدير، حيث:

$$B = 2.\delta_{\overline{X}_g}$$

أي أن متوسط المجتمع μ، وذلك عند مستوى ثقة $(1-\alpha)$ يقع في المدى (مجال الثقة) الآتي:

$$\boxed{\mu = \overline{X}_g \mp B}$$

2.1. مجال الثقة للمجموع الكلي للمجتمع T

علما أن تقدير المجموع الكلي T هو $\widehat{T} = N.\overline{X}_g$، بتباين: $\delta^2_{\widehat{T}} = N^2.\delta^2_{\overline{X}_g}$. فإن إيجاد مجال الثقة للمجموع الكلي T باستخدام طريقة حد الخطأ في تقدير المجموع يكون كما يلي:

إذا كان لدينا: $B' = 2.\delta_{\widehat{T}} = 2.N.\delta_{\overline{X}_g}$

فإن المجموع الكلي T، وذلك عند مستوى ثقة $(1-\alpha)$ يقع في المدى (مجال الثقة) الآتي:

$$\boxed{T = \widehat{T} \mp B'}$$

نشير هنا إلى أنه في حالة عدم توفر قيمة حجم المجتمع الكلي N، فإنه يمكن تقدير المجموع الكلي T بالعبارة $\widehat{T} = G.\overline{T}_g$، حيث $\overline{T}_g = \frac{\sum_{i=1}^{g} T_i}{g}$. وفي هذه الحالة تصبح عبارة تباين $(\widehat{T})$ كما يلي: $\delta^2_{\widehat{T}} = G^2.\delta^2_{\overline{T}_g}$، حيث $\delta^2_{\overline{T}_g} = \left(\frac{G-g}{g.G}\right)\frac{\sum_{i=1}^{g}(T_i-\overline{T}_g)^2}{g-1}$.

[25] عدنان عباس حميدان، مرجع سبق ذكره، ص453.

وعليه فإن مجال الثقة للمجموع الكلي T، وذلك عند مستوى ثقة $(1-\alpha)$ باستخدام طريقة حد الخطأ في التقدير يصبح:

$$T = \widehat{T} \mp B'$$

حيث:

$$B' = 2.\delta_{\widehat{T}} = 2.G.\delta_{\overline{T}_g}$$

مثال 6-1:

لدراسة متوسط الدخل الشهري للأسرة الواحدة في إحدى البلديات المكونة من 60 حي سكني، وعلى اعتبار كل حي بمثابة عنقود، تقرر الاعتماد على المعاينة العشوائية العنقودية من مرحلة واحدة لأجل هذه الدراسة، فتم إختيار 03 أحياء بإستعمال المعاينة العشوائية البسيطة وإستجواب مختلف الأسر المكونة لهم، وفيما يلي جدول يلخص أهم البيانات التي تم جمعها لهذه الغاية:

العنقود i	عدد الأسر N_i	إجمالي الدخل الشهري للأسر (دج) $T_i = \sum_{j=1}^{N_i} X_{ij}$
1	120	51.125
2	100	43.560
3	80	33.750
المجموع	300	128.435

المطلوب:

1/ أوجد مجال الثقة لمتوسط الدخل الشهري الأسري في هذه البلدية μ بحيث لا يزيد حد الخطأ في التقدير عن $\left(B = 2.\delta_{\overline{X}_g}\right)$ وذلك في الحالتين:

أ. عدد الأسر بالبلدية $N = 5.800$.

ب. N غير معلومة.

2/ أوجد مجال الثقة لإجمالي المداخيل الشهرية للأسر في هذه البلدية T بحيث لا يزيد حد الخطأ في التقدير عن $\left(B = 2.\delta_{\overline{X}_g}\right)$ وذلك في الحالتين:

أ. عدد الأسر بالبلدية $N = 5.800$.

ب. N غير معلومة.

الحل:

1/ مجال ثقة μ:

$$\overline{X}_g = \frac{\sum_{i=1}^{3} T_i}{\sum_{i=1}^{3} N_i} = \frac{128435}{300} = 428{,}11.$$

أ. عدد الأسر بالبلدية $N = 5.800$:

$$\overline{N} = \frac{N}{G} = \frac{5800}{60} = 96{,}66.$$

$$\delta^2_{\overline{X}_g} = \left(\frac{G-g}{g.G.\overline{N}^2}\right)\frac{\sum_{i=1}^{g}(T_i - N_i\overline{X}_g)^2}{g-1} = \left(\frac{60-3}{3.(60).(96{,}66)^2}\right)\frac{871337{,}56}{3-1} = 14{,}76.$$

$$\delta_{\overline{X}_g} = \sqrt{14{,}76} = 3{,}84.$$

$$B = 2.\delta_{\overline{X}_g} = 2.(3{,}84) = 7{,}68.$$

وعليه فإن مجال ثقة μ هو:

$$\mu \in [428.11 - 7.68, 428.11 + 7.68]$$

$$\mu \in [420.43\ , 435.79]$$

أي أن قيمة μ لا يمكن أن تقل عن 420,43دج ولا أن تزيد عن 435,79دج.

ب. N غير معلومة:

$$\overline{n} = \frac{n}{g} = \frac{300}{3} = 100.$$

$$\delta^2_{\overline{X}_g} = \left(\frac{60-3}{3.(60).(100)^2}\right)\frac{871337{,}56}{3-1} = 13{,}79.$$

$$\delta_{\overline{X}_g} = \sqrt{14{,}76} = 3{,}71.$$

$$B = 2.\delta_{\overline{X}_g} = 2.(3{,}71) = 7{,}42.$$

وعليه فإن مجال ثقة μ هو:

$$\mu \in [428.11 - 7.42, 428.11 + 7.42]$$

$$\mu \in [420.69\ , 435.53]$$

أي أن قيمة μ لا يمكن أن تقل عن 420,69دج ولا أن تزيد عن 435,53دج.

2/ مجال ثقة T:

أ. عدد الأسر بالبلدية $N = 5.800$:

$$\widehat{T} = N.\overline{X}_g = 5800.(428{,}11) = 2.483.038$$

$$B' = 2.\delta_{\widehat{T}} = 2.N.\delta_{\overline{X}_g} = 2.(5800).(3{,}84) = 44.544$$

وعليه فإن مجال ثقة T هو:

$$T \in [2483038 - 44544, 2483038 + 44544]$$

$$T \in [2438494\ ,\ 2527582]$$

أي أن قيمة T لا يمكن أن تقل عن 2.438.494دج ولا أن تزيد عن 2.527.582دج.

ب. N غير معلومة:

$$\overline{T}_g = \frac{\sum_{i=1}^{g} T_i}{g} = \frac{128435}{3} = 42.811,66.$$

$$\widehat{T} = G.\overline{T}_g = 60.(42811,66) = 2.568.699,60.$$

$$\delta^2_{\overline{T}_g} = \left(\frac{G-g}{g.G}\right)\frac{\sum_{i=1}^{g}(T_i-\overline{T}_g)^2}{g-1} = \left(\frac{60-3}{3.(60)}\right)\frac{151.785.316,67}{3-1} = 24.032.675,14.$$

$$\delta_{\overline{T}_g} = \sqrt{24032675,14} = 4.902,31.$$

$$B' = 2.(60).(4902,31) = 588.277,20.$$

وعليه فإن مجال ثقة T هو:

$$T \in [2568699.60 - 588277.20\ ,\ 2568699.60 + 588277.20]$$

$$T \in [1980422.40\ ,\ 3156976.80]$$

أي أن قيمة T لا يمكن أن تقل عن 1.980.422.40دج ولا أن تزيد عن 3.156.976,80دج.

2. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ والمجموع الكلي للمجتمع T

يعتمد تحديد حجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها لتقدير μ وT على عدد العناقيد في مجتمع الدراسة والحجم النسبي لهذه العناقيد. ولنفرض أن حجم العناقيد وعددها في مجتمع الدراسة G قد تم تحديده سلفا فإن المطلوب هنا هو تحديد عدد العناقيد g اللازم سحبها من بين هذه العناقيد.

بفرض أن التباين الحقيقي لمجتمع الدراسة هو δ_g^2 فإن تباين الوسط الحسابي للمجتمع باستخدام المعاينة العنقودية من مرحلة واحدة يعطى بالعلاقة:

$$\delta^2_{\overline{X}_g} = \left(\frac{G-g}{g.G.\overline{N}^2}\right).\delta_g^2$$

في حالة عدم توفر قيمة التباين الحقيقي للمجتمع δ_g^2، يمكن أن نستعمل بدلها مقدرتها غير المتحيزة S_g^2 حيث:

$$S_g^2 = \frac{\sum_{i=1}^{g}(T_i - N_i\overline{X}_g)^2}{g-1}$$

1.2. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ

لتقدير حجم العينة g اللازم سحبها باستخدام المعاينة العشوائية العنقودية من مرحلة واحدة وذلك باستخدام طريقة حد الخطأ في تقدير μ حيث:

$$.B = 2.\delta_{\overline{X}_g}, D = \frac{\overline{N}^2.B^2}{4}$$

وبنفس الطريقة المستعملة في العينة العشوائية الطبقية نجد:

$$\boxed{g = \frac{G.\delta_g^2}{G.D + \delta_g^2}}$$

نشير هنا إلى أنه في حالة عدم توفر قيمة $\overline{N}$، فإنه يمكن أن نستخدم بدلها مقدرتها $\bar{n}$.

2.2. تقدير حجم العينة اللازم لتقدير المجموع الكلي T

لتقدير حجم العينة g اللازم سحبها باستخدام المعاينة العشوائية العنقودية من مرحلة واحدة وذلك باستخدام طريقة حد الخطأ في تقدير المجموع الكلي T حيث:

$$.B' = 2.\delta_{\widehat{T}} = 2.N.\delta_{\overline{X}_g}, D' = \frac{B'^2}{4G^2}$$

وبنفس الطريقة المستعملة سابقا في تقدير حجم العينة اللازم سحبها لتقدير μ نجد:

$$\boxed{g = \frac{G.\delta_g^2}{G.D' + \delta_g^2}}$$

مثال 6-2:

بالعودة لبيانات المثال 6-1، **المطلوب:**

- أوجد تقدير لحجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها لتقدير μ بحيث لا يزيد حد الخطأ في التقدير عن $B = 10$ وذلك في الحالتين:

أ. عدد الأسر بالبلدية $N = 5.800$.

ب. N غير معلومة.

الحل:

- حجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها:

نحسب أولا:

$$S_g^2 = \frac{\sum_{i=1}^{g}(T_i - N_i\overline{X}_g)^2}{g-1} = \frac{871337,56}{3-1} = 435.702,84.$$

أ. عدد الأسر بالبلدية $N = 5.800$:

$$B = 10,\ D = \frac{\overline{N}^2.B^2}{4} = \frac{(96,66)^2.(10)^2}{4} = 233.578,89.$$

$$g = \frac{G.S_g^2}{G.D+S_g^2} = \frac{60.(435702,84)}{60.(233578,89)+435702,84} = 1,81.$$

حجم العينة اللازم سحبها هي $g = 2$ عنقودين.

ب. N غير معلومة:

$$D = \frac{\overline{n}^2.B^2}{4} = \frac{(100)^2.(10)^2}{4} = 250.000,$$

$$g = \frac{60.(435702,84)}{60.(250000)+435702,84} = 1,69.$$

حجم العينة اللازم سحبها هي $g = 2$ عنقودين.

3. تقدير نسبة المفردات (P) التي تتصف بصفة معينة في المجتمع

لتقدير نسبة المفردات التي تتصف بصفة معينة باستخدام المعاينة العشوائية العنقودية من مرحلة واحدة P $(\widehat{P})$، نضع: $\widehat{P}_g = \frac{\sum_{i=1}^{g} r_i}{\sum_{i=1}^{g} N_i}$، حيث يعطى تباينها المقدر بالعلاقة:

$$\delta^2_{\widehat{P}_g} = \left(\frac{G-g}{g.G.\overline{N}^2}\right)\frac{\sum_{i=1}^{g}(r_i - N_i\widehat{P}_g)^2}{g-1}$$

نشير هنا إلى أنه في حالة عدم توفر قيمة $\overline{N}$، فإنه يمكن أن نستخدم بدلها مقدرتها $\overline{n}$.

1.3.مجال الثقة للنسبة P

يمكن إيجاد مجال الثقة للنسبة P بالنسبة للعينة العنقودية من مرحلة واحدة أيضا باستخدام طريقة حد الخطأ في التقدير، حيث:

$$B = 2.\delta_{\widehat{P}_g}$$

أي أن النسبة P، وذلك عند مستوى ثقة $(1-\alpha)$ تقع في المدى (مجال الثقة) الآتي:

$$\boxed{P = \widehat{P}_g \mp B}$$

2.3. تقدير حجم العينة اللازم سحبها لتقدير النسبة P

لتقدير حجم العينة g اللازم سحبها باستخدام المعاينة العشوائية العنقودية من مرحلة واحدة وذلك باستخدام طريقة حد الخطأ في تقدير النسبة P حيث:

$$.B = 2.\delta_{\widehat{P}_g},\ D = \frac{\overline{N}^2B^2}{4}$$

وبنفس الطريقة المستعملة سابقا في تقدير حجم العينة اللازم سحبها لتقدير μ و T نجد:

$$\boxed{g = \frac{G.\delta_g^2}{G.D+\delta_g^2}}$$

تجدر الاشارة إلى أنه في حالة عدم توفر قيمة δ_g^2، يمكن أن نستعمل بدلها مقدرتها غير المتحيزة S_g^2 حيث:

$$S_g^2 = \frac{\sum_{i=1}^{g}(r_i - N_i\widehat{P}_g)^2}{g-1}$$

نفس الشيء في حالة عدم توفر قيمة $\overline{N}$، فإنه يمكن أن نستخدم بدلها مقدرتها $\bar{n}$.

مثال 6-3:

بالعودة لبيانات المثال 6-1، وفي نفس سياق الدراسة تم إحصاء عدد الأسر التي يقل دخلها الشهري عن 300دج (r_i) في الأحياء السكنية الثلاث المكونة للعينة المختارة، فكانت النتائج كما يلي:

العنقود i	عدد الأسر N_i	r_i
1	120	60
2	100	42
3	80	36
المجموع	300	138

المطلوب:

1/ أوجد مجال الثقة لنسبة الأسر التي يقل دخلها الشهري عن 300دج P بحيث لا يتعدى حد الخطأ في التقدير عن $\left(B = 2.\delta_{\widehat{P}_g}\right)$ وذلك في الحالتين:

أ. عدد الأسر بالبلدية $N = 5.800$.

ب. N غير معلومة.

2/ أوجد تقدير لحجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها لتقدير P بحيث لا يزيد حد الخطأ في التقدير عن $B = 0.05$ وذلك في الحالتين:

أ. عدد الأسر بالبلدية $N = 5.800$.

ب. N غير معلومة.

الحل:

1/ مجال ثقة P:

$$\widehat{P}_g = \frac{\sum_{i=1}^{g} r_i}{\sum_{i=1}^{g} N_i} = \frac{138}{300} = 0{,}46.$$

أ. عدد الأسر بالبلدية $N = 5.800$:

$$\delta^2_{\hat{P}_g} = \left(\frac{G-g}{g.G.\bar{N}^2}\right)\frac{\sum_{i=1}^{g}(r_i - N_i\hat{P}_g)^2}{g-1} = \left(\frac{60-3}{3.(60).(96,66)^2}\right)\frac{39,68}{3-1} = 0,00067.$$

$$\delta_{\hat{P}_g} = \sqrt{0,00067} = 0,026.$$

$$B = 2.\delta_{\hat{P}_g} = 2.(0,026) = 0,052.$$

وعليه فإن مجال ثقة P هو:

$$P \in [0,46 - 0,052\ ,\ 0,46 + 0,052]$$

$$P \in [0.408\ ,\ 0.512]$$

أي أن قيمة P لا يمكن أن تقل عن 0,408 ولا أن تزيد عن 0,512.

ب. N غير معلومة.

$$\delta^2_{\hat{P}_g} = \left(\frac{60-3}{3.(60).(100)^2}\right)\frac{39,68}{3-1} = 0,00063.$$

$$\delta_{\hat{P}_g} = \sqrt{0,00067} = 0,025.$$

$$B = 2.\delta_{\hat{P}_g} = 2.(0,025) = 0,05.$$

وعليه فإن مجال ثقة P هو:

$$P \in [0,46 - 0,05\ ,\ 0,46 + 0,05]$$

$$P \in [0.41\ ,\ 0.51]$$

أي أن قيمة P لا يمكن أن تقل عن 0,41 ولا أن تزيد عن 0,51.

2/ حجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها:

نحسب أولا:

$$S^2_g = \frac{\sum_{i=1}^{g}(r_i - N_i\hat{P}_g)^2}{g-1} = \frac{39,68}{3-1} = 19,84.$$

أ. عدد الأسر بالبلدية $N = 5.800$:

$$B = 0,05.$$

$$D = \frac{\bar{N}^2.B^2}{4} = \frac{(96,66)^2.(0,05)^2}{4} = 5,84.$$

$$g = \frac{G.S^2_g}{G.D+S^2_g} = \frac{60.(19,84)}{60.(5,84)+19,84} = 3,21.$$

حجم العينة اللازم سحبها هي $g = 4$ عناقيد، وبالتالي يجب إضافة أو سحب عنقود آخر من المجتمع.

ب. N غير معلومة:

$$D = \frac{\bar{n}^2.B^2}{4} = \frac{(100)^2.(0,05)^2}{4} = 6,25.$$

$$g = \frac{G.S_g^2}{G.D+S_g^2} = \frac{60.(19,84)}{60.(6,25)+19,84} = 3,01.$$

حجم العينة اللازم سحبها هي $g = 3$ عناقيد.

سلسلة تمارين مقترحة حول المحور السادس:

تمرين 01:

أجب عن الأسئلة التالية:

- ✓ قدم تعريفا مختصرا للعينة العشوائية العنقودية متعددة المراحل.
- ✓ ماهي مفردات العينة المسحوبة في حالة العينة العنقودية من مرحلة واحدة ؟
- ✓ ماهي مفردات العينة المسحوبة في حالة العينة العنقودية من مرحلتين ؟

تمرين 02:

لأجل تخصيص غلاف مالي للمنحة المدرسية للسنة الدراسية المقبلة، قررت مديرية التربية لإحدى الولايات إجراء دراسة لتحديد عدد التلاميذ المتمدرسين الذين لا تزيد أعمارهم عن 18 سنة بهذه الولاية. ولهذا الغرض قامت باختيار 10 أحياء سكنية من بين 100 حي سكني موجود بهذه الولاية وذلك باستخدام أسلوب المعاينة العشوائية العنقودية من مرحلة واحدة وتم جمع البيانات الموضحة في الجدول التالي:

رقم الحي السكني i	عدد الأسر في كل حي سكني N_i	عدد التلاميذ المتمدرسين في كل حي سكني i $T_i = \sum_{j=1}^{N_i} X_{ij}$
1	1.200	3.750
2	800	2.500
3	1.200	3.625
4	600	1.915
5	1.400	4.135
6	800	2.375
7	800	2.260
8	1.000	3.085
9	1.400	4.315
10	800	2.455
المجموع	10.000	30.415

المطلوب:

1/ أوجد تقدير لمتوسط عدد التلاميذ المتمدرسين الذين لا تزيد أعمارهم عن 18 سنة في الأسرة الواحدة بهذه الولاية $(\overline{X}_g)$.

2/ أوجد تقدير لعدد التلاميذ المتمدرسين الذين لا تزيد أعمارهم عن 18 سنة بهذه الولاية $(\widehat{T})$.

3/ أوجد مجال الثقة لمتوسط عدد التلاميذ المتمدرسين الذين لا تزيد أعمارهم عن 18 سنة في الأسرة الواحدة بهذه الولاية μ وذلك عند مستوى ثقة 0.95.

4/ أوجد مجال الثقة لعدد التلاميذ المتمدرسين الذين لا تزيد أعمارهم عن 18 سنة بهذه الولاية T بمستوى ثقة 0.95 وذلك في الحالتين:

أ. عدد الأسر بالولاية $N = 92.500$.

ب. N غير معلومة.

5/ أوجد تقدير لحجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها لتقدير μ بحيث لا يزيد حد الخطأ في التقدير عن $B = 0.1$ وذلك في الحالتين:

أ. عدد الأسر بالولاية $N = 92.500$.

ب. N غير معلومة.

6/ علما أن عدد الأسر بالولاية هو $N = 92.500$، ما هو حجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها لتقدير T بحيث لا يزيد حد الخطأ في التقدير عن $(B' = 5.000)$ تلميذ متمدرس؟ من خلال النتيجة المتوصل إليها، ما الذي يجب على مديرية التربية القيام به من أجل الحصول على نتائج يمكن تعميمها على المجتمع الكلي؟

<u>تمرين 03:</u>

في نفس إطار الدراسة الموضحة في التمرين رقم 02 ولأجل توفير حافلات للنقل المدرسي للتلاميذ الذين يقيمون بعيدا عن المؤسسات التعليمية، وعلى اعتبار أن عدد التلاميذ المتمدرسين في ال10 أحياء السكنية المكونة للعينة المختارة بمثابة عينة جديدة، عمدت مديرية التربية أيضا للاستفسار لدى الأسر المستجوبة عن مدى حاجة أطفالهم للنقل المدرسي، فكانت الأجوبة كما يلي:

رقم الحي السكني i	عدد التلاميذ المتمدرسين في كل حي سكني N_i	عدد التلاميذ الذين هم في حاجة للنقل المدرسي في كل حي سكني r_i
1	3.750	1.239
2	2.500	1.629
3	3.625	1.196
4	1.915	1.273
5	4.135	1.377
6	2.375	1.581
7	2.260	1.506
8	3.085	1.028
9	4.315	1.437
10	2.455	1.623
المجموع	30.415	13.889

المطلوب:

1/ أوجد مجال الثقة لنسبة التلاميذ الذين هم في حاجة للنقل المدرسي بالولاية P بحيث لا يتعدى حد الخطأ في التقدير عن $\left(B = 2.\delta_{\widehat{P}_g}\right)$ وذلك في الحالتين:

أ. عدد التلاميذ المتمدرسين بالولاية $N = 300.000$.

ب. N غير معلومة.

2/ أوجد تقدير لحجم العينة العشوائية العنقودية من مرحلة واحدة اللازم سحبها لتقدير P بحيث لا يزيد حد الخطأ في التقدير عن $B = 0.1$ وذلك في الحالتين:

أ. عدد التلاميذ المتمدرسين بالولاية $N = 300.000$.

ب. N غير معلومة.

المحور السابع: العينة العشوائية العنقودية من مرحلتين

تعد المعاينة العشوائية العنقودية من مرحلتين إمتدادا للمعاينة العشوائية العنقودية من مرحلة واحدة، تسمح للباحث بتوفير مزيد من الجهد، الوقت والمال بالمقارنة مع المعاينة العنقودية من مرحلة واحدة. فيتم تقسيم المجتمعات الكبيرة (لا يتطلب نوع المعاينة هنا إعداد الإطار الإحصائي لجميع مفردات الدراسة) إلى وحدات تسمى بوحدات المعاينة الأولية أو العناقيد (grappes)، ثم يتم اختيار عينة عشوائية بسيطة من هذه الوحدات الأولية، ونقوم كذلك بتقسيم الوحدات المختارة إلى وحدات ثانوية ثم اختيار عينة عشوائية بسيطة منها إلى أن نصل إلى عينة الدراسة التي تحوي على المفردات التي سيتم منها إستنباط المعلومات المطلوبة.

إذا كان المجتمع مقسما لوحدات أولية (عناقيد) ولنفترض أن عددها G، ونقوم باختيار عينة عشوائية بسيطة من هذه العناقيد وليكن عددها g، حيث: $N = \sum_{i=1}^{G} N_i$ وهي تمثل حجم مجتمع الدراسة وN_i تمثل حجم العنقود المختار i.

1. مجال الثقة بالنسبة لمتوسط المجتمع μ والمجموع الكلي T

على إعتبار X_{ij} تمثل قيمة المفردة j في العينة التي تم إختيارها من العنقود i، الذي تم إختياره في المرحلة الأولى من المعاينة، و$\overline{X}_i$ الوسط الحسابي للعينة التي تم سحبها من العنقود i حيث:

$$\overline{X}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} X_{ij}$$

n_i : يمثل عدد المفردات المختارة عشوائيا من العنقود i من العناقيد التي تم إختيارها في العينة في المرحلة الثانية من المعاينة.

فإنه يمكن تقدير متوسط المجتمع μ $(\overline{X}_{cl2})$ بالنسبة للمعاينة العنقودية على مرحلتين بالعلاقة:

$$\overline{X}_{cl2} = \frac{G}{N.g} \sum_{i=1}^{g} N_i \overline{X}_i$$

حيث: $\overline{X}_{cl2}$ هي مقدرة غير متحيز لمتوسط المجتمع μ.

كما تعطى مقدرة تباين الوسط الحسابي $(\overline{X}_{cl2})$ بالعلاقة:

$$\hat{\delta}^2_{\overline{X}_{cl2}} = \left(\frac{G-g}{g.G.\overline{N}^2}\right) S_b^2 + \frac{1}{g.G.\overline{N}^2} \sum_{i=1}^{g} N_i \left(\frac{N_i - n_i}{n_i}\right) S_i^2$$

حيث $\overline{\mathrm{N}} = \frac{\mathrm{N}}{\mathrm{G}}$ وتمثل متوسط عدد المفردات في العنقود الواحد من مجموع العناقيد المكونة للمجتمع الكلي. نشير هنا إلى أنه في حالة عدم توفر قيمة $\overline{\mathrm{N}}$، فإنه يمكن إستخدام مقدرتها $\bar{\mathrm{n}} = \frac{\mathrm{n}}{\mathrm{g}}$ في العلاقة السابقة.[26]

كما يعرف S_{b}^2 على أنها التباين بين العناقيد (variance inter-grappes)، ويعطى بالعلاقة:

$$S_{\mathrm{b}}^2 = \frac{\sum_{\mathrm{i}=1}^{\mathrm{g}} (\mathrm{N_i}\overline{\mathrm{X}}_{\mathrm{i}} - \overline{\mathrm{N}}\overline{\mathrm{X}}_{\mathrm{cl2}})^2}{g-1}$$

ويعرف S_{i}^2 على أنها التباين داخل العنقود i (variance intra-grappes)، ويعطى بالعلاقة:

$$S_{\mathrm{i}}^2 = \frac{1}{n_i - 1} \sum_{\mathrm{j}=1}^{n_i} \left(\mathrm{X_{ij}} - \overline{\mathrm{X}}_{\mathrm{i}}\right)^2$$

نلاحظ إذن أن تباين الوسط الحسابي المقدر $\left(\delta_{\overline{\mathrm{X}}_{\mathrm{cl2}}}^2\right)$ يعتمد في إحتسابه على التباين بين العناقيد وكذلك التباين داخل العنقود.

1.1. مجال الثقة لمتوسط المجتمع μ

وكما رأينا سابقا بالنسبة لطرق المعاينة المختلفة، يمكن إيجاد مجال الثقة للمتوسط μ بالنسبة للعينة العنقودية من مرحلتين أيضا باستخدام طريقة حد الخطأ في التقدير، حيث:

$$\mathrm{B} = 2.\,\delta_{\overline{\mathrm{X}}_{\mathrm{cl2}}}$$

أي أن متوسط المجتمع μ، وذلك عند مستوى ثقة $(1-\alpha)$ يقع في المدى (مجال الثقة) الآتي:

$$\boxed{\mu = \overline{\mathrm{X}}_{\mathrm{cl2}} \mp \mathrm{B}}$$

مثال 7-1:

قررت شركة لصناعة الملابس الجاهزة تملك 25 ورشة خياطة إستخدام المعاينة العشوائية من مرحلتين لتقدير متوسط الأجور الشهرية للعاملين فيها. لهذا تم إختيار 5 ورشات منها عشوائيا، ومن ثم إختيار عدد من المستخدمين في كل ورشة من الورشات المختارة في المرحلة الأولى للمعاينة، وكانت البيانات المتعلقة بأجورهم الشهرية موضحة كما يلي من خلال الجدول أسفله.

رقم الورشة في العينة i	عدد العاملين في الورشة i، N_i	عدد العاملين في العينة من الورشة i، n_i	أجور العاملين في العينة من الورشة i، $\mathrm{X_{ij}}$
1	50	5	75،70،105،80،98
2	48	5	68،90،86،95،100
3	62	6	85،79،69،79،88،95
4	40	4	80،85،98،105
5	60	6	100،90،95،68،88،102

[26] عدنان عباس حميدان، مرجع سبق ذكره، ص453.

والمطلوب:

- إذا كان في المؤسسة 1200 عاملا، قدر متوسط الأجور الشهرية للعمال وضع حد على الخطأ في تقدير هذا المتوسط.

الحل:

$G = 25, g = 5, N = 1200.$

$\overline{X}_i$	S_i^2	$N_i\overline{X}_i$	$(N_i\overline{X}_i)^2$	$N_i\left(\frac{N_i - n_i}{n_i}\right)S_i^2$
85,6	229,3	4.280	18.318.400	103.185
87,8	150,2	4.214,4	17.761.167,4	62.002,56
82,5	79,9	5.115	26.163.225	46.235,47
92	132,66	3.680	13.542.400	47.760
90,5	151,1	5.430	29.484.900	81.594
المجموع		22.719,40	105.270.092,40	340.777,03

متوسط عدد العمال في الورشة الواحدة في المؤسسة: $\overline{N} = \frac{1200}{25} = 48$.

وعليه، يمكن تقدير متوسط الأجور الشهرية للعمال في المؤسسة $(\overline{X}_{cl2})$ كما يلي:

$$\overline{X}_{cl2} = \frac{25}{1200.(5)}.(22.719,40) = 94,66$$

كما يمكن تقدير التباين بين العناقيد S_b^2 كما يلي:

$$S_b^2 = \frac{\sum_{i=1}^{g}(N_i\overline{X}_i - \overline{N}\overline{X}_{cl2})^2}{g-1}$$

$$= \frac{1}{g-1}\left[\sum_{i=1}^{g}(N_i\overline{X}_i)^2 - 2\overline{N}\overline{X}_{cl2}\sum_{i=1}^{g}N_i\overline{X}_i + g.(\overline{N}\overline{X}_{cl2})^2\right]$$

$$= \frac{1}{5-1}\Big[(105.270.092,40) - 2.(48).(94,66).(22.719,40) + 5.\big((48).(94,66)\big)^2\Big] = 508.966,33$$

وتقدير تباين الوسط الحسابي للأجور الشهرية كما يلي:

$$\delta^2_{\overline{X}_{cl2}} = \left(\frac{25-5}{5.(25).(48)^2}\right).(508.966,33) + \frac{340.777,03}{5.(25).(48)^2} = 38,29$$

$$\delta_{\overline{X}_{cl2}} = \sqrt{38,29} = 6,19$$

ومنها فإن حد الخطأ في تقدير متوسط المجتمع μ هو:

$$B = 2.\delta_{\overline{X}_{cl2}} = 2.(6,19) = 12,38$$

أي أن متوسط الأجور الشهرية للعمال في هذه المؤسسة μ، وذلك عند مستوى ثقة قدره 95% يقع في المدى الآتي:

$$\mu = 94{,}66 \mp 12{,}38$$

أي

$$P(82{,}28 \leq \mu \leq 107{,}04) = 0{,}95$$

2.1. مجال الثقة للمجموع الكلي للمجتمع T

علما أن تقدير المجموع الكلي T هو: $\widehat{\mathrm{T}} = \mathrm{N}.\overline{\mathrm{X}}_{\mathrm{cl2}} = \frac{G}{g}\sum_{\mathrm{i}=1}^{\mathrm{g}} \mathrm{N_i}\overline{\mathrm{X}}_{\mathrm{i}}$،

بتباين: $\delta^2_{\widehat{\mathrm{T}}} = \mathrm{N}^2.\delta^2_{\overline{\mathrm{X}}_{\mathrm{cl2}}} = \left(\frac{\mathrm{G}-\mathrm{g}}{\mathrm{g}}\right).G.S^2_{\mathrm{b}} + \frac{G}{\mathrm{g}}\sum_{\mathrm{i}=1}^{\mathrm{g}} \mathrm{N_i}\left(\frac{\mathrm{N_i}-n_i}{n_i}\right)S^2_{\mathrm{i}}$، فإن إيجاد مجال الثقة للمجموع الكلي T باستخدام طريقة حد الخطأ في تقدير المجموع يكون كما يلي:

إذا كان لدينا: $\mathrm{B}' = 2.\delta_{\widehat{\mathrm{T}}} = 2.\mathrm{N}.\delta_{\overline{\mathrm{X}}_{\mathrm{cl2}}}$

فإن المجموع الكلي T، وذلك عند مستوى ثقة $(1-\alpha)$ يقع في المدى (مجال الثقة) الآتي:

$$\mathrm{T} = \widehat{\mathrm{T}} \mp \mathrm{B}'$$

مثال 7-2:

بالعودة إلى بيانات المثال رقم 7-1، أوجد تقدير المجموع الكلي لأجور العمال الشهرية في هذه المؤسسة وضع حد على الخطأ في تقديره.

الحل:

كنا قد أشرنا إلى أن تقدير المجموع الكلي T هو:

$$\widehat{\mathrm{T}} = 1200.(94{,}66) = 113.592$$

وأن تقدير تباينه هو:

$$\delta^2_{\widehat{\mathrm{T}}} = (1200)^2.(38{,}29) = 55.137.600$$

وعليه يكون حد الخطأ في تقدير المجموع الكلي للمجتمع T كما يلي:

$$\mathrm{B}' = 2.\sqrt{55.137.600} = 14.850{,}94$$

أي أن المجموع الكلي للأجور الشهرية للعمال في هذه المؤسسة، وذلك بإحتمال قدره 0,95 لا يقل عن 98.741,06 ولا يزيد عن 128.442,94، أي أن:

$$P(98.741{,}06 \leq \mu \leq 128.442{,}94) = 0{,}95$$

2. تقدير حجم العينة اللازم لتقدير متوسط المجتمع μ

يتم تحديد حجم العينة العشوائية العنقودية من مرحلتين اللازم سحبها لتقدير μ على مرحلتين، ويعتمد في المرحلة الأولى على تحديد عدد العناقيد g التي سيتم سحبها من بين G من العناقيد في مجتمع الدراسة. وفي المرحلة الثانية يتم تحديد حجم كل عينة $\mathrm{n_i}$ سيتم سحبها من عدد العناقيد g التي تم إختيارها في المرحلة الأولى من المعاينة.

يلاحظ أن تحديد قيم g و n_i يعتمد على عاملين[27]:

- التباين بين العناقيد المختلفة في المجتمع.
- التباين داخل كل عنقود من العناقيد في المجتمع.

وبشكل عام إذا كانت العناقيد غير متجانسة والتباين بينها كبير ولكن التباين بين مفرداتها قليل، فإننا نقوم بسحب عدد أكبر من العناقيد g وعدد أقل من المفردات من كل عنقود من العناقيد المختارة n_i. وعلى العكس من ذلك إذا كان التباين بين العناقيد صغير ولكن التباين بين مفردات كل منها كبير فإننا نختار عدد أقل من العناقيد g وعدد أكبر من المفردات من كل عنقود من العناقيد المختارة n_i.

ولتبسيط عملية التحليل، نفترض أنه لدينا مجتمع أحجام عناقيده متساوية، أي أن:

$$N_1 = N_2 \dots\dots\dots\dots\dots = N_G = \overline{N}$$

وأن أحجام العينات التي سيتم سحبها من كل عنقود من هذه العناقيد هي الأخرى متساوية، أي أن:

$$n_1 = n_2 \dots\dots\dots\dots\dots = n_g = n$$

فإن من الممكن (إذا أهملنا معامل التصحيح) تقدير متوسط المجتمع μ بإستخدام المعاينة العشوائية العنقودية من مرحلتين كما يلي:

$$\overline{X}_{cl2} = \frac{1}{g}\sum_{i=1}^{g} \overline{X}_i$$

$$\hat{\delta}^2_{\overline{X}_{cl2}} = \frac{\delta_b^2}{g} + \frac{\delta_w^2}{n.g}$$

حيث:

δ_b^2 هو التباين بين متوسطات العناقيد.

δ_w^2 هو التباين بين المفردات داخل العناقيد.

ولتقدير قيم g و n_i فإننا نسعى إلى تقليل تكاليف المعاينة إلى أقل حد ممكن لأجل قيمة معينة للتباين $\hat{\delta}^2_{\overline{X}_{cl2}}$، أو تقليل قيمة هذا الأخير إلى أقل حد ممكن لنفس الحد من التكاليف.

على إعتبار أن تكلفة معاينة العنقود الواحد من عناقيد المجتمع هي c_1، وتكلفة معاينة كل مفردة من أي عنقود هي c_2، وأن التكلفة الإجمالية للمعاينة هي C، فإن دالة التكاليف للمعاينة تعطى بالعلاقة:

[27] فتحي أحمد عاروري، مرجع سبق ذكره، ص292.

$$C = g.c_1 + g.n.c_2$$

وبإستخدام طريقة لاغرانج نجد أن قيمة n التي تجعل من التباين $\hat{\delta}^2_{\overline{X}_{cl2}}$ نهاية صغرى لنفس التكاليف الكلية للمعاينة C، حيث:

$$n = \sqrt{\frac{c_1 \delta_w^2}{c_2 \delta_b^2}}$$

وبعد تحديد عدد المفردات التي يجب سحبها من كل عنقود من العناقيد التي سيتم سحبها في المرحلة الأولى من المعاينة يمكن تحديد عدد العناقيد g التي يلزم سحبها في المرحلة الأولى من المعاينة إذا كان التباين $\hat{\delta}^2_{\overline{X}_{cl2}}$ ثابتا، وذلك بإستخدام عبارته السابقة الذكر.

حيث عبارة g تعطى بالعلاقة :

$$g = \frac{\delta_b^2}{\hat{\delta}^2_{\overline{X}_{cl2}}} + \frac{\delta_w^2}{n.\hat{\delta}^2_{\overline{X}_{cl2}}}$$

وكما أشرنا سابقا، فإن عدد المفردات الواجب سحبها من كل عنقود من العناقيد المختارة في المرحلة الأولى للمعاينة n يتناسب طرديا مع التباين داخل العناقيد δ_w^2 وعكسيا مع التباين بين العناقيد δ_b^2.

ومن الناحية العملية فإن قيم δ_w^2 وδ_b^2 تكون في العادة مجهولة ولذلك يمكن للباحث أن يحصل عليها من دراسات سابقة أو من دراسات أولية أو دراسة إستطلاعية، حيث يمكن تقديرها كما يلي:

$$S_w^2 = \frac{1}{g}\sum_{i=1}^{g} S_i^2 = \hat{\delta}_w^2$$

$$S_b^2 = S^2 - \frac{S_w^2}{n} = \hat{\delta}_b^2$$

$$S^2 = \frac{1}{g-1}\sum_{i=1}^{g} (\overline{X}_i - \overline{X}_{cl2})^2 = \hat{\delta}^2$$

$$S^2 = S_b^2 + \frac{S_w^2}{n}$$

وعليه فإننا نقوم بتقدير δ_{w}^2 ب$\mathrm{S}_{\mathrm{w}}^2$ ثم بتقدير δ^2 بS^2 وδ_{b}^2 ب$\mathrm{S}_{\mathrm{b}}^2$.

مثال 7-3:

ترغب شركة لتجارة التجزئة تملك 30 فرعا، يعمل في كل فرع منها 20 موظفا، في تقدير متوسط الأجر الشهري الذي يتقاضاه العامل الواحد، بحيث لا يزيد التباين $\hat{\delta}_{\bar{\mathrm{X}}_{\mathrm{cl2}}}^2$ عن 12 دينار، ما هو حجم العينة اللازم سحبها لهذه الدراسة بإستخدام المعاينة العشوائية العنقودية من مرحلتين علما أن هذه الشركة كانت قد أجرت دراسة مماثلة العام الماضي على عينة من 05 فروع سحبت من كل منها 07 موظفين وكانت النتائج كالآتي:

$S^2 = 20, \mathrm{S}_{\mathrm{w}}^2 = 105, c_1 = 11, c_2 = 2.$

الحل:

على إعتبار نتائج دراسة العام الماضي بمثابة دراسة أولية:

$G = 30, S^2 = 20, \mathrm{S}_{\mathrm{w}}^2 = 105, c_1 = 11, c_2 = 2.$

لدينا:

$$S_{\mathrm{b}}^2 = S^2 - \frac{S_{\mathrm{w}}^2}{n} = 20 - \frac{105}{7} = 20 - 15 = 5$$

وكذلك:

$$n = \sqrt{\frac{c_1 \delta_{\mathrm{w}}^2}{c_2 \delta_{\mathrm{b}}^2}} = \sqrt{\frac{11.(105)}{2.(5)}} = 10{,}74$$

$$n = 11$$

أي أن عدد العاملين الذي يجب إختياره من كل عنقود من العناقيد المختارة هو 11 عامل. تحديد قيمة n يسمح لنا بتحديد عدد العناقيد الواجب سحبها في المرحلة الأولى من المعاينة كما يلي:

$$g = \frac{\delta_{\mathrm{b}}^2}{\hat{\delta}_{\bar{\mathrm{X}}_{\mathrm{cl2}}}^2} + \frac{\delta_{\mathrm{w}}^2}{\mathrm{n}.\hat{\delta}_{\bar{\mathrm{X}}_{\mathrm{cl2}}}^2} = \frac{5}{12} + \frac{105}{11.(12)} = 1{,}21$$

$$g = 2$$

عدد العناقيد الواجب سحبها في المرحلة الأولى من المعاينة هو 2.

سلسلة تمارين مقترحة حول المحور السابع:

تمرين 01:

أجب عن الأسئلة التالية:

- ✓ قدم تعريفا مختصرا للعينة العشوائية العنقودية متعددة المراحل.
- ✓ ماهي مفردات العينة المسحوبة في حالة العينة العنقودية من مرحلتين ؟

تمرين 02:

قررت إحدى الثانويات دراسة ظاهرة غياب التلاميذ فيها. وكان عدد الشعب المدرسة بها 20 شعبة. فإختارت الإدارة إستخدام المعاينة العشوائية من مرحلتين حيث تم سحب 04 شعب منها بصورة عشوائية كمرحلة أولى، ثم إختيار عدد من التلاميذ في المرحلة الثانية من تلك الشعب المختارة، وكانت البيانات المتعلقة بغياباتهم خلال الموسم الدراسي محل البحث موضحة كما يلي من خلال الجدول أسفله.

رقم الشعبة في العينة i	عدد التلاميذ في الشعبة i، N_i	عدد التلاميذ المختارين في العينة من الشعبة i، n_i	عدد مرات غياب تلاميذ العينة المختارة X_{ij}
1	40	4	3،2،3،4
2	38	5	4،4،5،3،1
3	42	5	5،3،4،1،2
4	36	4	1،6،3،4

والمطلوب:

- إذا عدد التلاميذ في هذه الثانوية 800، قدر متوسط عدد مرات غياب كل تلميذ وضع حد على الخطأ في تقدير هذا المتوسط.
- أوجد تقدير المجموع الكلي لغيابات التلاميذ في هذه الثانوية وضع حد على الخطأ في تقديره.

تمرين 03:

ترغب شركة لصناعة الملابس الجاهزة تملك 40 ورشة خياطة يعمل في كل واحدة منها 25 موظفا، في تقدير متوسط الأجور الإضافية التي يتقاضاها كل عامل أسبوعيا، بحيث لا يزيد التباين $\hat{\delta}^2_{\bar{X}_{cl2}}$ عن 04 دنانير، ماهو حجم العينة اللازم سحبها لهذه الدراسة بإستخدام المعاينة العشوائية العنقودية من مرحلتين علما أن هذه الشركة كانت قد أجرت دراسة مماثلة العام الماضي على عينة من 05 ورشات وسحبت من كل منها 8 موظفين وكانت النتائج كالآتي:

$S^2 = 16, S_w^2 = 112, c_1 = 10, c_2 = 2.$

قائمة المراجع

المراجع باللغة العربية:

- ✓ إبراهيم علي، "مدخل في تقنيات المعاينة الإحصائية"، مجلس التخطيط القطري، الدوحة 2007.
- ✓ ثائر فيصل شاهر، "اختبار الفرضيات الاحصائية"، دار الحامد للنشر والتوزيع، عمان 2013.
- ✓ جابر أحمد بسيوني، "الاحصاء العام"، دار الوفاء لدنيا الطباعة والنشر، الاسكندرية 2013.
- ✓ حسين علوان مطلق، "جمع البيانات وطرق المعاينة"، العبيكان للنشر، الظهران 2010.
- ✓ فتحي أحمد عاروري، "المعاينة الاحصائية - طرقها واستخداماتها"، الأكاديميون للنشر والتوزيع، عمان 2013.
- ✓ عدنان عباس حميدان، "تقنيات المعاينة الاحصائية في الدراسات السوقية والتسويقية"، دار الإعصار العلمي للنشر والتوزيع، عمان 2015.

المراجع باللغة الأجنبية:

- ✓ David R. ANDERSON, Jeffrey D. CAMM, James J. COCHRAN, Dennis J. SWEENEY, Thomas A. WILLIAMS, " Statistiques pour l'économie et la gestion", Édition DE BOECK, Paris 2015.
- ✓ Éric DOR, "Économétrie", Édition PEARSON EDUCATION, Paris 2004.
- ✓ G.BAILLARGEON, J.RAINVILLE, "Statistique appliquée : tests statistiques, régression et corrélation", Édition Trois-Rivières, Québec 1979.
- ✓ Pascal ARDILLY, "Les techniques de sondage", Édition TECHNIP, Paris 2006.
- ✓ Régis BOURBONNAIS et Michel TERRAZA, "Analyse des séries temporelles en économie", Édition PUF, Paris 1998.
- ✓ Régis BOURBONNAIS, "Économétrie, Manuel et exercices corrigés - 6éme édition", Édition DUNOD, Paris 2005.

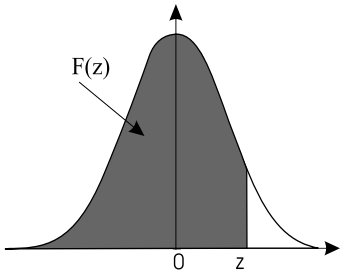
- ✓ Sandrine LARDIC et Valérie MIGNON, "Économétrie des séries temporelles macroéconomiques et financières", Édition ECONOMICA, Paris 2002.
- ✓ Yves TILLÉ, ″Théorie des sondages : échantillonnage et estimation en population finie : cours et exercices avec solutions″, Édition DUNOD, Paris 2011.

الجداول الإحصائية

1. جدول التوزيع الطبيعي
2. جدول توزيع ستيودنت
3. جدول توزيع كي دو

Fonction de répartition de la loi normale centrée réduite

(probabilité F(z) de trouver une valeur inférieure à z)



z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	0,5199	0,5239	0,5279	0,5319	0,5359
0,1	0,5398	0,5438	0,5478	0,5517	0,5557	0,5596	0,5636	0,5675	0,5714	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	0,5948	0,5987	0,6026	0,6064	0,6103	0,6141
0,3	0,6179	0,6217	0,6255	0,6293	0,6331	0,6368	0,6406	0,6443	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	0,7939	0,7967	0,7995	0,8023	0,8051	0,8078	0,8106	0,8133
0,9	0,8159	0,8186	0,8212	0,8238	0,8264	0,8289	0,8315	0,8340	0,8365	0,8389
1,0	0,8413	0,8438	0,8461	0,8485	0,8508	0,8531	0,8554	0,8577	0,8599	0,8621
1,1	0,8643	0,8665	0,8686	0,8708	0,8729	0,8749	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
1,4	0,9192	0,9207	0,9222	0,9236	0,9251	0,9265	0,9279	0,9292	0,9306	0,9319
1,5	0,9332	0,9345	0,9357	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	0,9505	0,9515	0,9525	0,9535	0,9545
1,7	0,9554	0,9564	0,9573	0,9582	0,9591	0,9599	0,9608	0,9616	0,9625	0,9633
1,8	0,9641	0,9649	0,9656	0,9664	0,9671	0,9678	0,9686	0,9693	0,9699	0,9706
1,9	0,9713	0,9719	0,9726	0,9732	0,9738	0,9744	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
2,3	0,9893	0,9896	0,9898	0,9901	0,9904	0,9906	0,9909	0,9911	0,9913	0,9916
2,4	0,9918	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	0,9941	0,9943	0,9945	0,9946	0,9948	0,9949	0,9951	0,9952
2,6	0,9953	0,9955	0,9956	0,9957	0,9959	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	0,9971	0,9972	0,9973	0,9974
2,8	0,9974	0,9975	0,9976	0,9977	0,9977	0,9978	0,9979	0,9979	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986

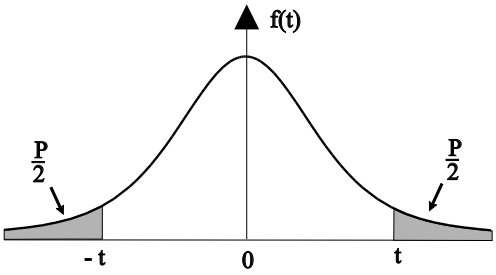
Table pour les grandes valeurs de z

z	3,0	3,1	3,2	3,3	3,4	3,5	3,6	3,7	3,8	3,9
F(z)	0,998650	0,999032	0,999313	0,999517	0,999663	0,999767	0,999841	0,999892	0,999928	0,999952
z	4,0	4,1	4,2	4,3	4,4	4,5	4,6	4,7	4,8	4,9
F(z)	0,999968	0,999979	0,999987	0,999991	0,999995	0,999997	0,999998	0,999999	0,999999	1,000000

Nota. La table donne F(z) pour z positif. Pour z négatif, il faut prendre le complément à l'unité de la valeur lue dans la table. Exemple : F(-1,37) = 1 - F(1,37) =1 - 0,9147 = 0,0853.

Table de la loi de Student

Valeurs de **T** ayant la probabilité **P** d'être dépassées en valeur absolue

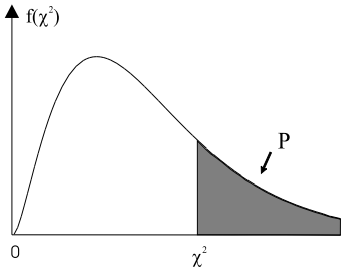


ν	P =0,90	0,80	0,70	0,60	0,50	0,40	0,30	0,20	0,10	0,05	0,02	0,01
1	0,158	0,325	0,510	0,727	1,000	1,376	1,963	3,078	6,314	12,706	31,821	63,657
2	0,142	0,289	0,445	0,617	0,816	1,061	1,386	1,886	2,920	4,303	6,965	9,925
3	0,137	0,277	0,424	0,584	0,765	0,978	1,250	1,638	2,353	3,182	4,541	5,841
4	0,134	0,271	0,414	0,569	0,741	0,941	1,190	1,533	2,132	2,776	3,747	4,604
5	0,132	0,267	0,408	0,559	0,727	0,920	1,156	1,476	2,015	2,571	3,365	4,032
6	0,131	0,265	0,404	0,553	0,718	0,906	1,134	1,440	1,943	2,447	3,143	3,707
7	0,130	0,263	0,402	0,549	0,711	0,896	1,119	1,415	1,895	2,365	2,998	3,499
8	0,130	0,262	0,399	0,546	0,706	0,889	1,108	1,397	1,860	2,306	2,896	3,355
9	0,129	0,261	0,398	0,543	0,703	0,883	1,100	1,383	1,833	2,262	2,821	3,250
10	0,129	0,260	0,397	0,542	0,700	0,879	1,093	1,372	1,812	2,228	2,764	3,169
11	0,129	0,260	0,396	0,540	0,697	0,876	1,088	1,363	1,796	2,201	2,718	3,106
12	0,128	0,260	0,395	0,539	0,695	0,873	1,083	1,356	1,782	2,179	2,681	3,055
13	0,128	0,259	0,394	0,538	0,694	0,870	1,079	1,350	1,771	2,160	2,650	3,012
14	0,128	0,258	0,393	0,537	0,692	0,868	1,076	1,345	1,761	2,145	2,624	2,977
15	0,128	0,258	0,393	0,536	0,691	0,866	1,074	1,341	1,753	2,131	2,602	2,947
16	0,128	0,258	0,392	0,535	0,690	0,865	1,071	1,337	1,746	2,120	2,583	2,921
17	0,128	0,257	0,392	0,534	0,689	0,863	1,069	1,333	1,740	2,110	2,567	2,898
18	0,127	0,257	0,392	0,534	0,688	0,862	1,067	1,330	1,734	2,101	2,552	2,878
19	0,127	0,257	0,391	0,533	0,688	0,861	1,066	1,328	1,729	2,093	2,539	2,861
20	0,127	0,257	0,391	0,533	0,687	0,860	1,064	1,325	1,725	2,086	2,528	2,845
21	0,127	0,257	0,391	0,532	0,686	0,859	1,063	1,323	1,721	2,080	2,518	2,831
22	0,127	0,256	0,390	0,532	0,686	0,858	1,061	1,321	1,717	2,074	2,508	2,819
23	0,127	0,256	0,390	0,532	0,685	0,858	1,060	1,319	1,714	2,069	2,500	2,807
24	0,127	0,256	0,390	0,531	0,685	0,857	1,059	1,318	1,711	2,064	2,492	2,797
25	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,316	1,708	2,060	2,485	2,787
26	0,127	0,256	0,390	0,531	0,684	0,856	1,058	1,315	1,706	2,056	2,479	2,779
27	0,127	0,256	0,389	0,531	0,684	0,855	1,057	1,314	1,703	2,052	2,473	2,771
28	0,127	0,256	0,389	0,530	0,683	0,855	1,056	1,313	1,701	2,048	2,467	2,763
29	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,311	1,699	2,045	2,462	2,756
30	0,127	0,256	0,389	0,530	0,683	0,854	1,055	1,310	1,697	2,042	2,457	2,750
∞	0,126	0,253	0,385	0,524	0,674	0,842	1,036	1,282	1,645	1,96	2,326	2,576

Nota. ν est le nombre de degrés de liberté.
Le quantile d'ordre $1 - \frac{\alpha}{2}$ se lit dans la colonne $P = \alpha$.
Le quantile d'ordre $1 - \alpha$ se lit dans la colonne $P = 2\alpha$.

Table de la loi du Khi-deux

Valeurs de χ^2 ayant la probabilité **P** d'être dépassées



ν	P = 0,995	0,99	0,975	0,95	0,90	0,10	0,05	0,025	0,01	0,005
1	0,00004	0,0002	0,001	0,0039	0,0158	2,706	3,841	5,024	6,635	7,879
2	0,010	0,020	0,051	0,103	0,211	4,605	5,991	7,378	9,210	10,597
3	0,072	0,115	0,216	0,352	0,584	6,251	7,815	9,348	11,345	12,838
4	0,207	0,297	0,484	0,711	1,064	7,779	9,488	11,143	13,277	14,860
5	0,412	0,554	0,831	1,145	1,610	9,236	11,070	12,833	15,086	16,750
6	0,676	0,872	1,237	1,635	2,204	10,645	12,592	14,449	16,812	18,548
7	0,989	1,239	1,690	2,167	2,833	12,017	14,067	16,013	18,475	20,278
8	1,344	1,646	2,180	2,733	3,490	13,362	15,507	17,535	20,090	21,955
9	1,735	2,088	2,700	3,325	4,168	14,684	16,919	19,023	21,666	23,589
10	2,156	2,558	3,247	3,940	4,865	15,987	18,307	20,483	23,209	25,188
11	2,603	3,053	3,816	4,575	5,578	17,275	19,675	21,920	24,725	26,757
12	3,074	3,571	4,404	5,226	6,304	18,549	21,026	23,337	26,217	28,300
13	3,565	4,107	5,009	5,892	7,042	19,812	22,362	24,736	27,688	29,819
14	4,075	4,660	5,629	6,571	7,790	21,064	23,685	26,119	29,141	31,319
15	4,601	5,229	6,262	7,261	8,547	22,307	24,996	27,488	30,578	32,801
16	5,142	5,812	6,908	7,962	9,312	23,542	26,296	28,845	32,000	34,267
17	5,697	6,408	7,564	8,672	10,085	24,769	27,587	30,191	33,409	35,718
18	6,265	7,015	8,231	9,39	10,865	25,989	28,869	31,526	34,805	37,156
19	6,844	7,633	8,907	10,117	11,651	27,204	30,144	32,852	36,191	38,582
20	7,434	8,260	9,591	10,851	12,443	28,412	31,410	34,170	37,566	39,997
21	8,034	8,897	10,283	11,591	13,240	29,615	32,671	35,479	38,932	41,401
22	8,643	9,542	10,982	12,338	14,041	30,813	33,924	36,781	40,289	42,796
23	9,260	10,196	11,689	13,091	14,848	32,007	35,172	38,076	41,638	44,181
24	9,886	10,856	12,401	13,848	15,659	33,196	36,415	39,364	42,980	45,559
25	10,520	11,524	13,120	14,611	16,473	34,382	37,652	40,646	44,314	46,928
26	11,160	12,198	13,844	15,379	17,292	35,563	38,885	41,923	45,642	48,290
27	11,808	12,879	14,573	16,151	18,114	36,741	40,113	43,195	46,963	49,645
28	12,461	13,565	15,308	16,928	18,939	37,916	41,337	44,461	48,278	50,993
29	13,121	14,256	16,047	17,708	19,768	39,087	42,557	45,722	49,588	52,336
30	13,787	14,953	16,791	18,493	20,599	40,256	43,773	46,979	50,892	53,672

Nota. ν est le nombre de degrés de liberté.
Pour $\nu > 30$, on peut admettre que la quantité $\sqrt{2\chi^2} - \sqrt{2\nu - 1}$ suit la loi normale centrée réduite.

الدكتور طلال عمراني

من مواليد بلدية عين وسارة (ولاية الجلفة)

دكتوراه علوم في الاقتصاد والاحصاء التطبيقي

أستاذ محاضر – أ – بجامعة الجلفة – الجزائر -

محتوى المطبوعة:

هذه المطبوعة هي مجموعة محاضرات خاصة بمقياس تقنيات المعاينة الاحصائية، وهي موجهة لطلبة السنة الثانية ماستر، شعبة العلوم التجارية، تخصص تسويق مصرفي، وهذا من أجل معرفة مختلف طرق جمع مفرادت عينة الدراسة، والتي يعتمد إختيارها على مدى درجة تمثيل العينة لخصائص المجتمع الأصل. كما يتطرق العمل لمختلف طرق تقدير خصائص وميزات المجتمع بالإعتماد على العينة المختارة التي ستشملها عملية الاستقراء.

إن إعداد هذه المطبوعة جاء لتحقيق جملة من الأهداف ونوجزها فيما يلي:

1. تدعيم مكتبة كلية العلوم الاقتصادية والعلوم التجارية وعلوم التسيير، وهذا من أجل تغطية العجز في هذا المجال.
2. تعريف الطالب بمختلف طرق المعاينة الاحصائية.
3. معرفة الطالب لطرق تقدير مختلف معلمات المجتمع الإحصائي من خلال عينة الدراسة.