جامعة الملك فيصل

عمادة التعلم الإلكتروني والتعليم عن بعد

اسم المقرر

التحليل الإحصائي

أستاذ المقرر

د/ محمد زايــد

المحاضرة (7)

مقدمة في المعاينة

تقديم :-

- تهتم نظرية المعاينة بدراسة العلاقة بين المجتمع والعينات المسحوبة منه فيما يسمى بالاستدلال الإحصائي statistical inference
- يعتبر الاستدلال الإحصائي من أهم الأدوات المساعدة على اتخاذ القرارات في الاقتصاد والأعمال وباقي العلوم، ويعني الاستدلال الإحصائي تقدير قيمة أو قيم غير معلومة تخص مجتمع الدراسة اعتمادا على بيانات عينة مأخوذة من هذا المجتمع ، وبمعنى آخر هو تعميم نتائج العينة على مجتمع الدراسة بكامله.
- تتمثل أدوات الاستدلال أو الاستنتاج الإحصائي بشكل أساسي في كل من التقدير واختبار الفرضيات.
- لكى يكون التقدير واختبار الفرضيات سليماً، ينبغي أن يبنى على عينة ممثلة للمجتمع يتم اختيارها وفقا لطبيعة الهدف من الدراسة.

المجتمع والعينة:-

المجتمع Population

لأي مجموعة من المفردات تشترك في صفة أو صفات محددة وتكون موضوع دراسة أو بحث ، فإن هذه المجموعة يطلق عليها إحصائيا مجتمع الدراسة أو اختصاراً المجتمع Population.

والمجتمع قد يكون مجموعة ما من البشر أو أشجار أنواع معينة من الفاكهة أو الحيوانات الزراعية أو إنتاج دولة ما لسلع معينه خلال فترة زمنية محدده...الخ.

والمجتمع قد يكون محدوداً إذا كان يمكن حصر عدد أفراده مثل سكان مدينة ما أو طلاب مرحلة تعليمية معينة ، وقد يكون المجتمع غير محدود (لانهائي) إذا كان لا يمكن حصر عدد أفراده مثل النجوم والكواكب أو الكائنات الحية بمياه المحيطات والأنهار.

المجتمع والعينة:-

وعند دراسة صفة ما أو صفات معينه لمجتمع ما فإن البيانات الإحصائية عن تلك الصفة أو الصفات تجمع بأحد أسلوبين:-

أولاً: أسلوب الحصر الشامل (census): وفيه تجمع البيانات عن كل مفرده من مفردات المجتمع، وهذا الأسلوب لا يتبع عادة إلا في حالة التعدادات التي تجريها الدول وتدعمها بإمكانيات ضخمه مثل تعدادات السكان والتعدادات الصناعية والتعدادات الزراعية.

ثانيا: أسلوب المعاينة (Sampling): وفيه يتم جمع البيانات عن جزء من مفردات المجتمع يختار بطريقة أو بأخرى ويطلق عليه عينة (Sample) ثم بعد ذلك يتم تعميم نتائج الدراسة على المجتمع بأكمله.

المعالم Parameters والإحصاءات Statistics:-

اعتاد البعض على معاملة القيم التي يحصل عليها من العينة وكأنها قيم مجتمعها، وهذا خطأ فادح.

فالمقاييس الإحصائية التي تحسب من بيانات مجتمع الدراسة بأكمله يطلق عليها معالم أو مؤشرات المجتمع (Parameters of population) ، أما المقاييس الإحصائية التي تحسب من بيانات عينة مسحوبة من مجتمع الدراسة فيطلق عليها إحصاءات (Statistics).

وللتفرقة بين المعالم والإحصاءات يجب أن نرمز لكل منها برموز تختلف عن رموز الأخرى، على سبيل المثال يرمز للمتوسط الحسابي للمجتمع بالرمز μ بينما يرمز للمتوسط الحسابي للعينة بالرمز $\bar{x}$ ، أيضاً يرمز للانحراف المعياري للمجتمع بالرمز σ بينما يرمز للانحراف المعياري للعينة بالرمز s ... وهكذا

بعض مزايا أسلوب المعاينة:-

يتميز أسلوب المعاينة عن أسلوب الحصر الشامل بمزايا عديدة منها:

1. يؤدي استخدام العينات العشوائية إلى خفض تكاليف الدراسات الميدانية بسبب صغر حجم العينة بالنسبة إلى حجم المجتمع وهو ما يؤدي إلى تخفيض الأعباء الإدارية والفنية التي تتطلبها أي دراسة ميدانية.

2. يتحقق وفر واضح في الوقت الذي ينفق في دراسة ميدانية على أساس عينة بدلاً من الحصر الشامل وتتضح أهمية الوقت عندما نقوم بدراسة ظاهرة تتغير بمرور الوقت، فتكون البيانات المجموعة والنتائج وقت ظهورها غير مطابقة لواقع المجتمع وتصبح النتائج ذات قيمه محدودة بعد أن فقدت عنصر المطابقة مع واقع الظاهرة وتوزيعها الحالي في المجتمع.

بعض مزايا أسلوب المعاينة:

3. في المجتمعات غير المحدودة (اللانهائية) مثل مجتمع الكائنات الحية في البحار والمحيطات لا يمكن أن تتم الدراسة على أساس الحصر الشامل ولكن لابد وأن تتم الدراسة بأسلوب المعاينة.

4. أيضاً هناك بعض الاختبارات لابد وأن تتم بأسلوب المعاينة لأن إجراء مثل هذه الاختبارات على أساس الحصر الشامل يؤدي إلى تلف المادة المختبرة أو هلاكها.. فاختبار صلاحية شحنة من المفرقعات مثلاً لابد وأن يتم على أساس العينة وبالمثل تحليل دم المرضى يتم على أساس عينة.

أقسام العينات:-

هناك عدة طرق لأخذ العينات من المجتمع لاستخدامها في الاستدلال الإحصائي. وبشكل عام تنقسم العينات عادة إلى قسمين رئيسين وهما عينات عشوائية وعينات غير عشوائية.

1. العينات العشوائية:

وهي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفردة فيها ، حيث يتم الاختيار باستخدام أساليب معينة تلعب الصدفة خلالها الدور الأول في اختيار المفردة ولكن بشرط أن يتحقق لجميع المفردات احتمال ثابت ومحدد للاختيار. والعينات العشوائية إذا ما تم اختيارها بالطريقة العلمية السليمة والمناسبة يمكن أن تكفل درجه عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة.

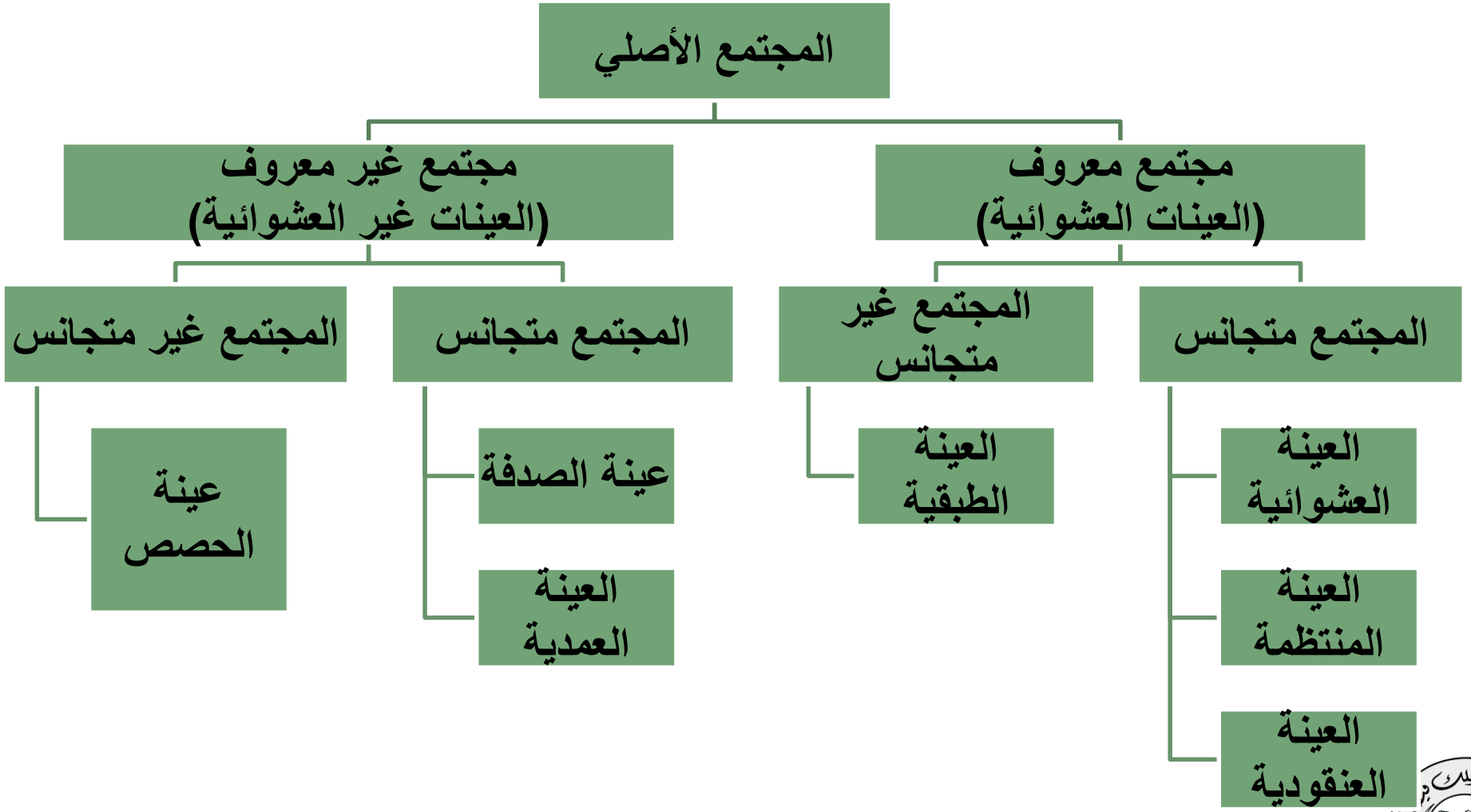
أقسام العينات:-

2. العينات غير العشوائية:

وهي تلك العينات التي لا تكفل لجميع مفردات المجتمع احتمال ثابت ومحدد للاختيار، وغالباً يتدخل الباحث في عملية الاختيار بصورة أو بأخرى.

وفيما يلي استعراض لأهم أنواع العينات العشوائية والعينات غير العشوائية.

أقسام العينات:-



أ- العينات الاحتمالية:-

العينة العشوائية البسيطة	جميع عناصر المجتمع لها نفس الفرصة في الظهور في العينة
العينة الطبقية	يقسم المجتمع إلى طبقتين على الأقل ثم نختار عينة عشوائية من كل طبقة تتناسب وحجمها.
العينة المنتظمة	ترتب مفردات المجتمع ثم نختار نقطة بداية للمعاينة ونحدد بعد ثابت من هذه النقطة لاختيار باقي النفاط تباعا.
العينة العنقودية	يقسم المجتمع إلى مساحات أو أجزاء ثم نختار عشوائيا بعض هذه المساحات، ثم نختار جميع أو بعض عناصرها بالعينة.

ب- العينات غير الاحتمالية:

عينة الصدفة	يتم اختيارها عن طريق الصدفة
العينة العمدية (القصدية)	يتم اختيار أفراد العينة تحت شروط معينة لتحقيق الهدف من التجربة
عينة الحصص	يقسم المجتمع إلى أجزاء ثم نختار العينة من كل جزء من أجزاء المجتمع وفقا للنسب المحددة

اختيار العينة:

- يطلق على المصدر الذي تؤخذ منه العينة إطار المعاينة وهو حصر شامل لجميع مفردات مجتمع الدراسة.
- يمكن أن يقسم إطار المعاينة إلي أقسام تسهل عملية الاختيار يطلق على كل قسم منها وحدة معاينة.
- يؤثر حجم مجتمع الدراسة في اختيار مفردات العينة. إذا كان حجم المجتمع صغيراً جداً من الممكن عدم الحصول على عدد كاف من المفردات أما إذا كان حجم المجتمع كبيراً ـ وهذا هو المتوقع دائماً ـ تكون المشكلة في كيفية اختيار العينة.
- كلما كثرت الشروط التي يجب توفرها في مفردات العينة كلما صعب الحصول على العدد المطلوب.

اختيار العينة:

خطوات اختيار العينة:

أ- تحديد أهداف البحث.

ب- تحديد كل من المجتمع الأصل الذي تختار منه العينة ووحدة المعاينة (إطار المعاينة).

ج- إعداد قائمة بالمجتمع الأصل بحيث يكون لكل مفردة فيه رقم أو تمييز محدد.

د- اختيار الطريقة المتبعة في المعاينة وتحديد حجم العينة.

هـ - الحصول على عينة مناسبة.

تحديد حجم العينة:

يعتبر تحديد حجم العينة المناسب من أهم قرارات الباحث للحصول على بيانات تزوده بمعلومات يمكن الاعتماد عليها لتعميم النتائج. ويتوقف حجم العينة الواجب دراسته على تفاعل بعض العوامل من أهمها:

1. حجم المجتمع.
   (كلما زاد حجم مجتمع الدراسة، يزيد حجم العينة المطلوب).
2. مدى التباين في خصائص المجتمع المراد دراسته.
   (كلما زاد التباين، يزيد حجم العينة المطلوب).
3. مدى الخطأ الذي يُسمح به في نتائج العينة كتقديرات لخصائص المجتمع.
   (كلما قل مدى الخطأ الذي يمكن السماح به، زاد حجم العينة)
4. درجة الثقة التي نود أن نتمتع بها في تحقق السمات السابقة.
   (كلما زادت درجة الثقة المطلوبة، زاد حجم العينة اللازم)

تحديد حجم العينة:

توجد عدة طرق لتحديد حجم العينة المناسب بما يلائم منهج البحث، وحجم وطبيعة المجتمع، وكذلك الغرض من العينة.

وتعتمد طرق حساب حجم العينة على معادلات أو صيغ رياضية مختلفة ، كما توجد مجموعة من الجداول التي صممت وفقا لبعض هذه الصيغ تحدد الحجم المناسب للعينة من المجتمع الأصلي، ويستفيد منها الباحثون الذين لا يميلون إلى الأسلوب الرياضي.

وسنهتم هنا فقط بصيغتين مبسطتين لحساب حجم العينة تستخدم إحداهما إذا كان الغرض من العينة هو تقدير متوسط المجتمع وتستخدم الأخرى إذا كان الغرض من العينة هو تقدير نسبة وجود صفة أو خاصية معينة بين مفردات المجتمع.

تحديد حجم العينة:

(1) حجم العينة المناسب لتقدير متوسط المجتمع

$$n = \left(\frac{z\ \sigma}{d}\right)^2$$

n : حجم العينة المطلوب.

z : قيمة مستخرجة من جدول التوزيع الطبيعي المعياري تتحدد بناء على درجة الثقة المطلوبة. (درجة ثقة 95% : z = 1.96 ، درجة ثقة 99% : z = 2.58)

σ : الانحراف المعياري لقيم المجتمع. (إذا لم يكن معلوما من دراسات سابقة ، يمكن تقديره بربع المدى المتوقع لقيم المجتمع)

d : خطأ التقدير الذي يحدده الباحث وفقا لطبيعة دراسته.

تحديد حجم العينة:

مثال (1)

حدد حجم العينة المناسب لتقدير متوسط أعمار طلاب كلية إدارة الأعمال إذا كنا نرغب في ألا يزيد خطأ التقدير عن 2 سنة وبدرجة ثقة 95% علما بأن الانحراف المعياري لأعمار الطلاب من واقع الخبرة السابقة هو 4 سنوات.

الحل

حيث أن درجة الثقة المطلوبة هي 95 % ، فإن z = 1.96

$$n = \left(\frac{z\ \sigma}{d}\right)^2 = \left(\frac{(1.96)\ (4)}{2}\right)^2 = 15.37 = 16$$

أي أن حجم العينة المطلوب يساوي 16 طالبا.

تحديد حجم العينة:

مثال (2)

حدد حجم العينة المناسب لتقدير متوسط دخل العامل في صناعة معينة بالمملكة العربية السعودية إذا كنا نرغب في ألا يزيد خطأ التقدير عن 50 ريال وبدرجة ثقة 95% علما بأن الانحراف المعياري لدخول العمال في هذا المجال من واقع الخبرة السابقة هو 200 ريال.

الحل

حيث أن درجة الثقة المطلوبة هي 95% ، فإن z = 1.96

$$n = \left(\frac{z\ \sigma}{d}\right)^2 = \left(\frac{(1.96)\ (200)}{50}\right)^2 = 61.47 = 62$$

أي أن حجم العينة المطلوب يساوي 62 عاملا.

تحديد حجم العينة:

مثال (3)

حدد حجم العينة المناسب لتقدير متوسط أطوال طلاب جامعة الملك فيصل إذا كنا نرغب في ألا يزيد خطأ التقدير عن 10 سم وبدرجة ثقة 99%.

الحل

حيث أن درجة الثقة المطلوبة هي 99% ، فإن z = 2.58

حيث أن الانحراف المعياري لأطوال الطلاب غير معلوم ، نقوم بتقدير المدى الممكن للطول ، فمثلا نتوقع أن تكون الأطوال بين 120 سم و 200 سم ، ثم يقدر الانحراف المعياري بربع المدى ، ويكون: $\sigma = \frac{200-120}{4} = 20$

$$n = \left(\frac{z\ \sigma}{d}\right)^2 = \left(\frac{(2.58)\ (20)}{10}\right)^2 = 26.63 = 27$$

أي أن حجم العينة المطلوب يساوي 27 طالبا.

تحديد حجم العينة:

(2) حجم العينة المناسب لتقدير النسبة فى المجتمع

$$n = \left(\frac{z}{d}\right)^2 p(1-p)$$

حيث:

n : حجم العينة المطلوب.

z : القيمة من جدول التوزيع الطبيعي المعياري (z = 2.58 or z = 1.96)

p : نسبة الصفة المطلوب تقديرها في المجتمع من واقع الخبرة السابقة أو الدراسات المشابهة. (إذا كانت هذه النسبة غير معلومة ، نضع p = 0.5)

d : خطأ التقدير الذي يحدده الباحث وفقا لطبيعة دراسته.

تحديد حجم العينة:

مثال (4)

حدد حجم العينة المناسب لتقدير نسبة البطالة في إحدى المدن إذا كنا نرغب في ألا يزيد خطأ التقدير عن 3% وبدرجة ثقة 99% علما بأن هذه النسبة من واقع الخبرة السابقة كانت 30%.

الحل

حيث أن درجة الثقة المطلوبة هي 95% ، فإن z = 2.58

$$n = \left(\frac{z}{d}\right)^2 p(1-p) = \left(\frac{2.58}{0.03}\right)^2 (0.3)(0.7) = 1553.16 = 1554$$

أي أن حجم العينة المطلوب يساوي 1554 فردا.

تحديد حجم العينة:

مثال (5)

حدد حجم العينة المناسب لتقدير نسبة المدخنين من بين طلاب جامعة الملك فيصل إذا كنا نرغب في ألا يزيد خطأ التقدير عن 5% وبدرجة ثقة 95%.

الحل

حيث أن درجة الثقة المطلوبة هي 95% ، فإن z = 1.96

حيث أن النسبة p غير معلومة من مصادر سابقة ، p = 0.5

$$n = \left(\frac{z}{d}\right)^2 p(1-p) = \left(\frac{1.96}{0.05}\right)^2 (0.5)(0.5) = 384.16$$

أي أن حجم العينة المطلوب يساوي 385 طالبا.

أخطاء البيانات الإحصائية: -

تتعرض البيانات الإحصائية التي يتم جمعها إلى نوعين من الأخطاء:

1. **خطأ التمييز أو التحيز:** وهو ذلك الخطأ الناتج عن مصادر متعددة، منها أخطاء في تصميم البحث أو التجربة أو أخطاء فنيه أثناء جمع البيانات أو خلال العمليات الحسابية التي تتم على البيانات المتجمعة.
2. **خطأ المعاينة العشوائية أو خطأ الصدفة:** وهو الخطأ الناتج عن فروق الصدفة بين مفردات المجتمع التي دخلت العينة وبين تلك المفردات التي لم تشأ الصدفة أن تدخل العينة.

خطأ التمييز أوالتحيز:

الأسباب:

- الاختيار غير العشوائي للعينة: تعتمد بعض طرق الاختيار للعينة على خاصية معينة كالاعتماد على دليل الهاتف (عند دراسة الدخل والإنفاق).
- التحيز المقصود (تعمد إدخال بعض الوحدات)
- استبدال وحدة بوحدة أخرى غير مدرجة ضمن الإطار العام للدراسة.

تفادي خطأ التحيز:

- اختيار جميع وحدات العينة عشوائياً باستخدام إحدى طرق الاختيار العشوائي
- عدم استبدال أية وحدة تم اختيارها بوحدة أخرى
- تدريب الباحثين بشكل جيد على جمع البيانات والتقيد بالتعليمات

خطأ المعاينة العشوائية Random Sampling Error

عند اختيار العينة العشوائية هناك خطأ ينتج عن الاختلاف أو التشتت بين قيم الوحدات التي تتكون منها العينة وتلك الوحدات التي كان لها فرصة أن تدخل في العينة ، وهذا الخطأ يسمى بخطأ المعاينة العشوائي أو خطأ الصدفة.

ويمكن تفادي أو تقليل خطأ المعاينة العشوائي عن طريق:

- زيادة حجم العينة (في حدوت التكلفة والوقت المتاح للمعاينة)
- تحري الدقة في اختيار أسلوب المعاينة المناسب (كالأسلوب الطبقي أو العينة المنتظمة...الخ) ، مما يسهم في تقليل الاختلاف بين الوحدات المختارة في العينة والوحدات التي كان يمكن أن يشملها الاختيار.

تمت بحمد الله

جميع الحقوق محفوظة لجامعة الملك فيصل ©
All Rights Reserved for KFU ©