

المحاضرة الأولى

تعريف المجموعة :-

يمكن تعريف المجموعة على أنها عدد من العناصر بينها صفات مشتركة تكتب بين حاصرتين { } و تسمى بأحد الحروف الهجائية الكبيرة
A , B , C ,

و من الأمثلة على المجموعات A = { 1,3,5,7,9,.....}

الأشياء التي تتكون منها المجموعة تسمى عناصر المجموعة و ترمز للعناصر بواسطة حروف صغير مثل :-

a , b , c ,

يستخدم الرمز ϵ "ينتمي إلى" لبيان عناصر المجموعة فمثلاً إذا كان العنصر a من ضمن عناصر المجموعة A فإننا نقول أن a ينتمي إلى المجموعة A و يكتب بالصورة $a \in A$

أما إذا كان a ليس عنصراً من عناصر المجموعة A فإننا نقول أن العنصر a لا ينتمي إلى المجموعة A و يكتب على الصورة $a \notin A$

طريقة كتابة المجموعات :

طريقة العد (سرد العناصر) :-

يتم فيها وضع جميع عناصر المجموعة ، أو جزء منها ، بين قوسى المجموعة { } بحيث يفصل بين كل عنصرين بعلامة فاصلة " , " :-
مثال :-

$$A = \{ 1,5,10,15 \}$$

$$B = \{ a , b , c , d \}$$

$$C = \{ 1 , 2 , 3 \dots \}$$

(و هي مجموعة منتظمة تسير بنفس الشكل ١ ٢ ٣ ٤ وهكذا)

$$A = \{ 1, 2, 3, \dots, 100 \}$$

(و هي مجموعة مغلقة و لكن المساحة لا تكفي لكتابة من ١ إلى ١٠٠ و سوف نستخدم النقاط للتعبير عن بعض العناصر)

طريقة القاعدة (الصفة المميزة) :-

ويتم فيها وصف المجموعة بذكر صفة يمكن بواسطتها تحديد عناصرها ، أي الصفة التي تحدد ارتباط عناصر المجموعة ، فمثلاً :

$$A = \{ x : \text{عدد زوجي } x \}$$

$$B = \{ x : \text{طالب بمقرر الاحصاء في الادارة } x \}$$

$$C = \{ x : \text{طالب بنظام التعليم عن بعد } x \}$$

$$D = \{ x : \text{عدد صحيح } -3 \leq x \leq 1 \}$$

$$X = \{ x : \text{عدد صحيح } 0 \leq x \leq 12 \}$$

أنواع المجموعات :-

❖ المجموعة الخالية :-

هي المجموعة التي لا تحتوي أي عنصر ويرمز لها بالرمز ϕ (فاي) أو { } .

❖ المجموعة المنتهية :-

المجموعة التي تكون عناصرها محدودة .

❖ المجموعة غير المنتهية :-

المجموعة التي تكون عناصرها غير محدودة (وهي المجموعة التي لا يمكن تحديد عناصرها بشكل دقيق)

❖ المجموعة الكلية :-

هي المجموعة التي تدرس جميع المجموعات باعتبارها مجموعات جزئية و يرمز لها بالرمز U .

❖ المجموعة الجزئية :-

تكون المجموعة A جزئية من المجموعة B إذا كانت جميع عناصر A موجودة في B و تكتب على الصورة :- $A \subset B$.

أمثلة :-

١- إذا كانت $A = \{2, 4, 6\}$ و $B = \{1, 2, 3, 4, 5, 6, 7, 8\}$ فإن $A \subset B$.

٢- المجموعة المكونة من جميع طلاب التعليم الالكتروني بجامعة الملك فيصل مجموعة جزئية من مجموعة طلاب هذه الجامعة .

❖ تساوي المجموعات :-

تكون المجموعتان A و B متساويتان إذا كانت :-

$$A \subseteq B , B \subseteq A \implies A = B$$

أما المجموعتان المتكافئتان فهما المجموعتان اللتان تتساويان في عدد عناصرها وتكتب على الصورة $A \equiv B$

\subseteq رمز يعني (مجموعة مساوية) يرمز لها (\equiv)

\equiv رمز يعني (مجموعة متكافئة)

العمليات على المجموعات :-

❖ الاتحاد :-

اتحاد المجموعتين A و B ($A \cup B$) هو مجموعة كل العناصر الموجودة في A **أو** في B أو في كليهما .

❖ التقاطع :-

تقاطع المجموعتين A و B ($A \cap B$) هو مجموعة كل العناصر الموجودة في A **و** في B معاً أي العناصر المشتركة بين A و B .

❖ المكملة أو المتممة :-

يقال أن \bar{A} مكملة المجموعة A إذا كانت تحتوي على جميع عناصر المجموعة الكلية U باستثناء عناصر A .

❖ الفرق :-

إذا كانت مجموعتان A ، B فإن $A - B$ يسمى بالفرق وهو مجموعة كل العناصر الموجودة في A وليست في B .

أشكال فن

VIN Figures

يمكن تمثيل المجموعات والعمليات المختلفة عليها من خلال استعمال اشكال هندسية تسمى أشكال فن وذلك وفق ما يلي:

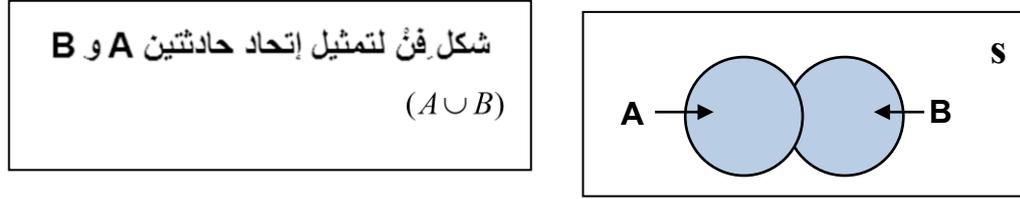
١- المجموعة الكلية:

تمثل المجموعات الكلية بمستطيل ويرمز لها بالرمز S

S

٢- اتحاد الحوادث Events Union :

لأي حادثتين A و B فإن الحادثة التي تتضمن كافة العناصر التي تنتمي إلى A أو إلى B أو كليهما معا يطلق عليها اتحاد حادثتين ويرمز لها $(A \cup B)$ أو (A أو B) والشكل التالي يوضح ذلك:



وبشكل عام لأي n حادثة $A_1, A_2, A_3, \dots, A_n$ فإن اتحاد هذه الحوادث هو :

$$\bigcup_{i=1}^n A_i = A_1 \cup A_2 \cup A_3 \dots \cup A_n$$

ويمكن القول أن $\bigcup_{i=1}^n A_i$ هو حدث يقع إذا وقع أحد هذه الحوادث A_i على الأقل وهو ما يطلق عليه جمع الأحداث فالإتحاد \cup يعني اتحاد المجموعتين A و B وهو مجموع العناصر الموجودة في كلتا المجموعتين دون تكرار العناصر.

مثال:

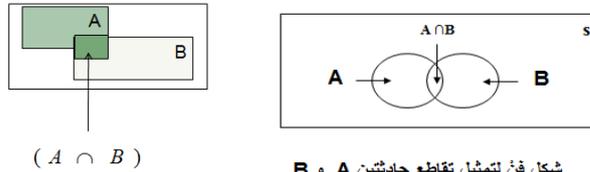
$$A = \{1, 2, -6, -7\}$$

$$B = \{-6, -7, -11\}$$

$$(A \cup B) = \{1, 2, -6, -7, -11\}$$

٣- تقاطع الحوادث Events Intersection :

لأي حادثتين A و B فإن الحادثة التي تتضمن كافة نقاط العينة التي تنتمي إلى A و B أو كليهما معا في نفس الوقت يطلق عليها تقاطع حادثتين ويرمز لها $(A \cap B)$ أو (A و B) وباستخدام أشكال فن يكون الجزء المحدد بـ A and B هو الذي يمثل تقاطع الحادثتين :



$$(A \cap B)$$

شكل فن لتمثيل تقاطع حادثتين A و B

وبشكل عام لأي n حادثة $A_1, A_2, A_3, \dots, A_n$ فإن تقاطع هذه الحوادث هو :

$$\bigcap_{i=1}^n A_i = A_1 \cap A_2 \cap A_3 \dots \cap A_n$$

ويمكن القول أن $\bigcap_{i=1}^n A_i$ هو حدث يقع إذا فقط وقعت كل الحوادث A_i على الأقل وهو ما يطلق عليه ضرب الحوادث

تقاطع الحوادث Events Intersection :

فالتقاطع \cap إذا هو مجموعة العناصر المشتركة بين مجموعتين أو أكثر.

خواص العمليات الجبرية لتقاطع الحوادث:

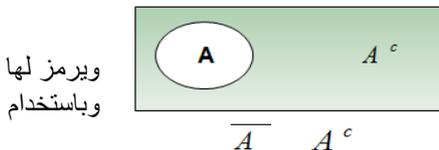
$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C) \quad \text{- إذا كانت A و B و C ثلاث حوادث فإن :}$$

ويعني ذلك توزيع التقاطع على الإتحاد.

$$(A \cap B) = (B \cap A) \quad \text{- وكذلك هناك خاصية التبديل والتي تعني أن}$$

٤- الحادثة المتممة Complementary Event :

لأي حادثة A فإن متممها هي الحادثة التي تتضمن كافة العناصر التي لا تنتمي إلى A بالرمز A^c أو \overline{A} وهو حدث يتألف من جميع عناصر Ω غير المنتمية إلى A أشكال فن فإن الجزء المظلل يمثل الحادثة المتممة :

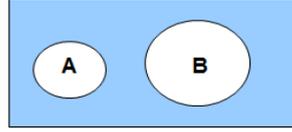


شكل فن لتمثيل مكملة الحادثة A

٥- الحوادث المتنافية Mutually Exclusive Events :

الحدثان A و B متنافيتان أو منفصلتان إذا كان تقاطعهما خالياً أي أن $A \cap B = \phi$

ويمكن القول أيضاً أن $A \cap A^c = \phi$ ، وباستخدام أشكال فن فإن الحادثتان المنفصلتان يتمثلان بالشكل التالي :



$$A \cap B = \phi$$

$$A \cap A^c = \phi$$

شكل فن لتمثيل حدثان متنافيتان A و B

بعض العلاقات المهمة :-

$$A \cup A^c = S$$

$$\overline{B \cup A} = \overline{B} \cap \overline{A}$$

$$A \cap A^c = \phi$$

$$\overline{A \cup B} = \overline{A} \cap \overline{B}$$

$$\overline{\overline{S}} = S$$

إذا كانت $A \subset B$ فإن:

$$\overline{\overline{\phi}} = \phi$$

$$A = A \cap B$$

$$A \cup S = S$$

$$B = A \cup B$$

$$A \cap S = A$$

$$\overline{\overline{B}} = B$$

$$A \cap \phi = \phi$$

\cap = و
 \cup = أو

أمثلة وتمارين

يراد شراء ثلاثة أنواع من اللحوم من جزر معين، فإذا رمزنا للحم الدجاج بـ A ولحم الضأن بـ B ، ولحم العجل بـ C فإن :

- توفر أنواع اللحوم الثلاثة يعني توفر لحم A و B و C أي بمعنى : $A \cap B \cap C$

- عدم توفر أي نوع من اللحوم يعني عدم توفر A و B و C أو كلها أي بمعنى : $\overline{A \cap B \cap C}$

- توفر نوع واحد من اللحوم على الأقل هو توفر A أو B أو C أو كلها أي بمعنى : $A \cup B \cup C$

- توفر نوع A فقط يعني : $A \cap \overline{B} \cap \overline{C}$

- توفر نوع واحد من اللحم يعني إما توفر A وعدم توفر النوعين الآخرين أو توفر B وعدم توفر النوعين الآخرين ، أو توفر C وعدم توفر النوعين الآخرين أي بمعنى :

$$(A \cap \overline{B} \cap \overline{C}) \cup (\overline{A} \cap B \cap \overline{C}) \cup (\overline{A} \cap \overline{B} \cap C)$$

المحاضرة الثانية

تعريف الاحتمالات:

يمكن تعريف الاحتمالات بطرق عديدة غير أن أبسطها

"هو مقياس لامكانية وقوع حدث (Event) معين"

وتستعمل كلمة احتمالات بشكل دائم في حياتنا اليومية مثل:

- احتمال أن ينجح الطالب في مقرر دراسي.
- احتمال نزول المطر اليوم

وفي أحيان أخرى تستخدم كلمة احتمالات ككلمة مرادفة لبعض الكلمات الأخرى مثل: ممكن، غالباً، مؤكد، مستحيل ...

وقد ارتبط المفهوم التقليدي للاحتمال بالألعاب الحظ لمدة طويلة، وتختلف درجة إمكانية تحقق أي حادثة من شخص إلى آخر حسب خبرته والمعلومات المتوفرة لديه عن الحادثة

وقد تطور علم الاحتمالات تطوراً كبيراً وسريعاً وأصبح أساساً لعلم الإحصاء وبحوث العمليات وغيرها.

١- التجربة العشوائية Random Experiment :

هي تلك التجربة التي تكون جميع نتائجها معلومة مسبقاً ولكن لا يمكن التنبؤ بحدوث أي من هذه النتائج بصفة مؤكدة مثلاً:

- رمي حجر نرد مرة واحدة يعتبر تجربة عشوائية، حيث نعلم جميع قيم نتائج التجربة وهي ١ أو ٢ أو ٣ أو ٤ أو ٥ أو ٦ ولكن لا يمكن تحديد أي الأرقام يظهر إلى الأعلى بصورة مؤكدة قبل إجراء التجربة.
- رمي عملة معدنية مرة واحدة أو عدد من المرات يعتبر تجربة عشوائية معروف جميع نتائجها قبل أن تبدأ التجربة، ولكن لا يمكن الجزم بظهور أي منها في رمية معينة.
- المشاركة في سباق الخيل لحصان معين يعتبر تجربة عشوائية فهو إما أن يفوز أو يخسر أو يتعادل.

٢- الحادث والفراغ العيني:

فراغ العينة هو المجموعة الشاملة التي تحتوي على جميع النتائج الممكنة للتجربة العشوائية ويرمز له بالرمز Ω ويطلق عليه الحالات الممكنة .

افترض أننا نقوم بإجراء تجربة ما كرمي زهرة النرد مثلاً ونلاحظ كل النتائج الممكنة وهي ظهور أحد الأوجه الستة ١ أو ٢ أو ٣ أو ٤ أو ٥ أو ٦ ونفترض أننا مهتمون بظهور رقم فردي أي ١ أو ٣ أو ٥ من التجربة . وهكذا فإن عملية رمي الزهرة تسمى **تجربة** (Experiment) وظهور رقم فردي هو محل اهتمامنا يسمى **حادثاً** (Event) ومجموعة جميع الحالات الممكنة الظهور تسمى **بالفراغ العيني** :

(Sample Space) ويلاحظ أن الحادث قد يكون حالة أو أكثر من الفراغ العيني .

الحادثة هي مجموعة جزئية من فراغ العينة وتمثل مجموعة النتائج التي تحقق الحدث وتسمى أيضا الحالات المواتية Favorable Cases ، فمثلاً الحصول على رقم زوجي في تجربة رمي زهرة النرد مرة واحدة تكون الحادثة هي {٢ ، ٤ ، ٦} ، ويمكن أن تحتوي الحادثة على عنصر واحد أو أكثر.

٣- أنواع الحوادث :

أ- الحوادث المتنافية (Mutually Exclusive Events)

يقال عن الحادثين A و B أنهما متنافيان إذا استحال حدوثهما معاً. فمثلاً عند رمي عملة معدنية لا يمكن الحصول على وجهين في وقت واحد.

ب- الحوادث المستقلة (Independent Events)

يعتبر الحادثين A أو B حادثين مستقلين إذا كان وقوع إحداهما أو عدم وقوعه لا يؤثر في وقوع الآخر. فمثلاً عند رمي قطعة عملة واحدة مرتين متتاليتين فإن نتيجة الرمية الثانية لا تتأثر بنتيجة الأولى.

ج- الحوادث الشاملة (Exhaustive Events)

تسمى الحوادث A ، B ، C ... حوادث شاملة في تجربة ما إذا كان لابد من حدوث إحداها عند إجراء التجربة.

فمثلاً عند اختيار طالب من الجامعة لمعرفة حالته ما إذا كان مدخناً أو غير مدخن تعتبر هذه الحالات حوادث شاملة لأنه لا بد للفرد أن يكون له صفة واحدة من هذه الصفات. كذلك فإن الحصول على العدد ١ أو ٢ أو ٣ أو ٤ أو ٥ أو ٦ عند رمي حجر النرد تعتبر حوادث شاملة لأنه لا بد من حدوث إحداها.

جمع الاحتمالات

أ- في حالة كون الحوادث متنافية

إذا كانت الحوادث A_1, A_2, A_3, \dots حوادث متنافية بمعنى أن حدوث أحدها يؤدي إلى استحالة حدوث أي من الحوادث الأخرى وبالتالي فإن احتمال حدوث هذه الحوادث معاً يكون معدوماً فإن

احتمال وقوع أي حادث من الحوادث المتنافية يساوي مجموع احتمالات وقوع هذه الحوادث

فإذا كان A, B حادثين متنافيين كما في الشكل (١)

$$P(A \text{ أو } B) = P(A) + P(B)$$

يرمز أيضاً لاحتمال وقوع أحد الحادثين بالرمز

حوادث متنافية

$$P(A \cup B) = P(A) + P(B)$$

ب- في حالة كون الحوادث غير متنافية

عند عدم اشتراط تنافي الحادثين A و B يكون المقصود بالحدث (A أو B) وقوع A على انفراد أو وقوع B على انفراد أو وقوع الحادثين A و B معاً في وقت واحد كما يتضح من الشكل التالي

الاحتمال الشرطي Conditional Probability

- الاحتمال الشرطي

إذا كان لدينا الحادثين A_1, A_2 وكان $P(A_2)$ لا يساوي الصفر فإن الاحتمال الشرطي للحادث A_1 بشرط وقوع الحادث A_2 يعطي بالمعادلة التالية:

$$P(A_1 | A_2) = \frac{P(A_1 \cap A_2)}{P(A_2)}$$

أي أن الاحتمال الشرطي للحادث A_1 بشرط وقوع الحادث A_2 يساوي حاصل قسمة الاحتمال المركب لـ A_1, A_2 على احتمال الحادث A_2

ضرب الاحتمالات

إن احتمال حدوث حادثين مستقلين أو أكثر معاً يساوي حاصل ضرب احتمال حدوث كل واحد من هذه الحوادث ببعضها بعضاً

فمثلاً إذا كان لدينا صندوق به ١٠ كرات متماثلة منها ٦ بيضاء و ٤ سوداء وسحبنا كرة من الصندوق فإن احتمال أن تكون بيضاء $6/10$ واحتمال أن تكون سوداء $4/10$ ، فإذا أعدنا الكرة إلى الصندوق (ليصبح العدد مكتملاً كما كان) وسحبنا الكرة مرة أخرى فإن احتمال أن تكون بيضاء $6/10$ واحتمال أن تكون سوداء $4/10$ ، فتكرار العملية يؤدي إلى نفس الاحتمال.

ومن هذا نرى أن نتيجة السحب الأول لا تؤثر على نتيجة السحب الثاني وهذا ما يسمى سحب بإرجاع أو إحلال أو إعادة.

فإذا كان لدينا الحادثين المستقلين A_1 و A_2 فإن احتمال حدوثهما معاً هو:

$$P(A_1 \cap A_2) = P(A_1) P(A_2)$$

بمعنى أن احتمال وقوع حدثين مستقلين معاً يساوي حاصل ضرب احتمال وقوع أي منهما بمفرده في احتمال وقوع الحدث الآخر بمفرده وفي حالة التعميم لـ n فإن :

$$P(A_1 \cap A_2 \cap A_3 \cap \dots \cap A_n) = P(A_1) P(A_2) P(A_3) \dots P(A_n)$$

مما سبق يمكن القول أن الحوادث المعطاة تكون مستقلة عندما تبقى الاحتمالات ثابتة مثل الحوادث:

١- رمي قطع نقود (أو قطعة واحدة عدة مرات)

٢- رمي أحجار نرد (أو حجر نرد عدة مرات)

٣- السحب مع الإرجاع (أو الإعادة)

نظرية بايز (Bayes' Theorem)

إذا كانت A_1, A_2, \dots, A_n مجموعة أحداث متنافية وكانت احتمالات حدوثها

$P(A_1), P(A_2), \dots, P(A_n)$ وإذا كان هناك حدث B يحدث إذا حدث أي من الأحداث المتنافية أنظر الشكل بالأسفل ، فإن احتمال حدوث الحدث A_r بشرط حدوث B هو :

$$P(A_r|B) = \frac{P(A_r)P(B|A_r)}{\sum_{i=1}^n P(A_i)P(B|A_i)} \quad 1 \leq r \leq n$$

المحاضرة الثالثة

المتغير العشوائي Random Variable:

المتغير العشوائي هو الذي يأخذ قيما حقيقية مختلفة تعبر عن نتائج فراغ العينة، ومن ثم مجال هذا المتغير، يشمل كل القيم الممكنة له، ويكون لكل قيمة من القيم التي يأخذها المتغير احتمال معين .

وينقسم المتغير العشوائي إلى قسمين هما:

- المتغيرات العشوائية المنفصلة Discrete Random Variables
- المتغيرات العشوائية المتصلة (المستمرة) Continuous Random Variables

أولاً : المتغيرات العشوائية المتقطعة (المنفصلة):

المتغير العشوائي المنفصل هو الذي يأخذ قيم بينية، ومتباعدة، ويرمز للمتغير العشوائي بشكل عام بحرف من الحروف الأبجدية الكبيرة X, Y, Z, \dots ويرمز للقيم التي يأخذها المتغير بالحروف الأبجدية الصغيرة، x, y, z, \dots

فالمتغير العشوائي المنفصل هو كل قيمة من قيم المتغير العشوائي ، ومن أمثلة هذه المتغيرات:-

- عدد الأولاد الذكور في الأسرة المكونة من أربع أولاد X ، $X: \{x=0,1,2,3,4\}$.
- عدد العملاء الذين يتم إنهاء خدمتهم البنكية كل 10 دقائق Y ، $Y: \{y=0,1,2,3, \dots\}$.
- عدد مرات استخدام نوع معين من الأسمدة خلال الدورة الزراعية.
- عدد الوحدات التالفة من إنتاج مزرعة معينة تنتج 200 وحدة كل موسم.
- عدد الوحدات التي تستهلكها الأسرة من سلعة معينة خلال الشهر.

التوزيع الاحتمالي للمتغير العشوائي المنفصل:

التوزيع الاحتمالي، هو الذي يبين احتمالات حدوث القيم التي يمكن أخذها المتغير، والتي ترتبط باحتمالات النتائج الممكنة في فراغ العينة، وبمعنى آخر هو التكرار النسبي للقيم التي يمكن أن يأخذها المتغير.

فإذا كان المتغير العشوائي المنفصل X يأخذ القيم، $X : \{x = x_1, x_2, \dots, x_n\}$ ،

وكان $P(X = x_i) = f(x_i)$ هو احتمال أن المتغير العشوائي يأخذ القيمة x_i

، فإنه، يمكن تكوين جدول التوزيع الاحتمالي للمتغير العشوائي X ، وهو جدول مكون من عمودين، الأول به القيم الممكنة للمتغير :

$$X : \{x = x_1, x_2, \dots, x_n\}$$

، والثاني به القيم الاحتمالية لهذا المتغير ، أي أن: $P(X = x_i) = f(x_i)$

جدول التوزيع الاحتمالي للمتغير العشوائي المنفصل

x_i	$f(x_i)$
x_1	$f(x_1)$
x_2	$f(x_2)$
\vdots	\vdots
x_n	$f(x_n)$
Σ	1

وتسمى الدالة $f(x_i)$ بدالة الاحتمال

الوسط الحسابي والتباين للمتغير العشوائي المنفصل:

يرمز للوسط الحسابي للمتغير العشوائي بالرمز (ميو)، ويحسب بتطبيق المعادلة التالية:

$$\mu = \sum x_i f(x_i) \quad (3-8)$$

وأما التباين ويرمز له بالرمز σ^2 (سيجما^٢)، فيحسب بتطبيق المعادلة التالية:

$$\begin{aligned} \sigma^2 &= \sum (x_i - \mu)^2 f(x_i) \\ &= \sum x_i^2 f(x_i) - \mu^2 \end{aligned} \quad (4-8)$$

ثانياً : المتغيرات العشوائية المستمرة Continuous Random Variables

المتغير العشوائي المستمر هو الذي يأخذ قيماً متصلة، ويأخذ عدد لانهايتي من القيم الممكنة له داخل مجاله، فإذا كان x متغير عشوائي مستمر، ويقع في المدى (a,b) ، أي أن: $\{X = x : a < x < b\}$ ، فإن للمتغير X عدد لانهايتي من القيم تقع بين الحدين الأدنى والأعلى (a,b) ، ومن الأمثلة على المتغيرات الكمية المستمرة ما يلي:

- كمية الألبان التي تنتجها البقرة في اليوم باللتر: $\{X = x : 10 < x < 40\}$
- المساحة المزروعة بالأعلاف في المملكة بالآلاف هكتار $\{X = x : 1000 < x < 15000\}$
- فترة صلاحية حفظ الدجاج المبرد بالأيام، $\{X = x : 1 < x < 5\}$
- وزن الجسم بالكيلوجرام للأعمار من $(30-40)$ ، $\{X = x : 55 < x < 80\}$
- وهكذا الأمثلة على المتغير الكمي المستمر كثيرة.

• الوسط الحسابي والتباين للمتغير العشوائي المستمر:

• إذا كانت $f(x)$ هي دالة كثافة الاحتمال للمتغير العشوائي x ، $a < x < b$ فإن معادلة الوسط والتباين يمكن كتابتها كما يلي:

$$\mu = E(x) = \int_a^b x f(x) dx$$

$$\sigma^2 = E(x^2) - \mu^2 , E(x^2) = \int_a^b x^2 f(x) dx$$

(٨-١٣)

التوزيعات الاحتمالية للمتغيرات المنفصلة:-

أ- توزيع ذي الحدين:

يستخدم هذا التوزيع في الحالات التي يكون للظاهرة محل الدراسة نتيجتان فقط متنافيتان، النتيجة محل الاهتمام وتسمى بحالة النجاح، والأخرى تسمى بحالة الفشل، ومن أمثلة ذلك:

- عند إعطاء مريض نوع معين من الأدوية، لها نتيجتان:
(استجابة للدواء، أو عدم استجابة)
- عند فحص عبوة بداخلها نوع معين من الفاكهة، لها نتيجتان
(الوحدة إما أن تكون سليمة، أو تكون معيبة)
- عند إلقاء قطعة عملة، لها نتيجتان
(ظهور الوجه الذي يحمل الصورة، أو الوجه الذي يحمل الكتابة)
- نتيجة الطالب في الاختبار
(نجاح، رسوب)
- استخدام المزارع لبرنامج معين في الزراعة
(يستخدم، أو لا يستخدم)

شكل التوزيع الاحتمالي ثنائي الحدين:

إذا كررت محاولة من المرات، بحيث أن كل محاولة لها نتيجتان فقط متنافيتان هما:

- النتيجة محل الاهتمام " حالة نجاح " وتتم باحتمال ثابت في كل محاولة هو p
- النتيجة الأخرى " حالة فشل " وتتم باحتمال ثابت أيضا هو $q = 1 - p$

وبافتراض أن هذه المحاولات مستقلة، بمعنى أن نتيجة كل محاولة ليس لها علاقة بنتيجة المحاولة الأخرى، وإذا كان المتغير العشوائي X يعبر عن عدد حالات النجاح "عدد النتائج محل الاهتمام" في الـ n محاولة، فإن مدي المتغير العشوائي X والذي يعبر عن عدد حالات النجاح هو:

$$X : \{x = 0, 1, 2, \dots, n\}$$

إذا فتوزيع ذو الحدين هو أحد التوزيعات الاحتمالية المنفصلة، ويستخدم لإيجاد احتمال وقوع حدث معين (نجاح) عدداً من المرات مقداره X من بين n من المحاولات لنفس التجربة (ونرمز لهذا الاحتمال بالرمز $P(X)$ وذلك عندما تتحقق الشروط التالية :

- هناك ناتجان ممكنان فقط ومتنافيان لكل محاولة .
- المحاولات وعددها n مستقلة عن بعضها البعض .
- احتمال وقوع الحدث المعين في كل محاولة (النجاح) P ثابت ولا يتغير من محاولة لأخرى .

فبالتالي يمكن حساب الاحتمال من خلال المعادلة التالية :

$$P(X) = \frac{n!}{X!(n-X)!} P^X (1-P)^{n-X}$$

حيث $n!$ (وتقرأ " مضروب n ") = $1 \cdot 2 \cdot 3 \cdot \dots \cdot (n-2) \cdot (n-1) \cdot n \cdot 1 = 0!$.

ويكون متوسط توزيع ذي الحدين $\mu = np$

وانحراف المعياري $\sigma = \sqrt{np(1-p)}$

تحديد شكل التوزيع:

يتحدد شكل التوزيع ثنائي الحدين وفقاً لقيمة احتمال النجاح كما يلي:

- إذا كان $p = 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون **متماثل**.
- إذا كان $p < 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون **موجب الالتواء**.
- إذا كان $p > 0.5$ فإن التوزيع الاحتمالي ثنائي الحدين يكون **سالب الالتواء**.

ب - توزيع بواسون:-

توزيع بواسون هو توزيع احتمالي منفصل آخر يستخدم لتحديد احتمال وقوع عدد معين من النجاحات في وحدة الزمن، وذلك عندما تكون الأحداث أو "النجاحات" مستقلة عن بعضها البعض وعندما يبقى متوسط عدد النجاحات ثابتاً لوحدة الزمن . عندئذ :

$$P(x) = \frac{e^{-\mu} \mu^x}{x!}$$

$$, x = 0, 1, 2, \dots$$

حيث : x = العدد المعين من النجاحات.

$P(x)$ = احتمال عدد x من النجاحات.

e = أساس نظام اللوغاريتمات الطبيعي وتوجد في بعض الآلات الحاسبة،

وقيمتها هي: $e = 2.718$ تقريباً، ويمكن حساب قيمتها باستخدام

الآلة الحاسبة.

$x!$ = مضروب العدد x " ويساوي: $x! = x(x-1)(x-2) \dots 3 \times 2 \times 1$

μ = المتوسط

• يشترك توزيع بواسون من توزيع ذي الحدين عندما يكون :-

- عدد المحاولات n كبير جداً
- بينما يكون احتمال النجاح p صغير بحيث تبقى np قيمة ثابتة معتدلة
- يوصف متغيرات عشوائية متقطعة تعبر عن عدد كبير من الحوادث مثل:
 - عدد حوادث السيارات في الشهر داخل مدينة كبيرة
 - عدد الكرات الحمراء في عينة الدم

○ عدد الأخطاء المطبعية في الصفحات المختلفة للكتاب

○ عدد القطع التالفة في الإنتاج الكلي لسلعة معينة

إذا كان للمتغير X توزيع بواسون فإن

$$E(X) = \lambda \text{ التوقع}$$

$$\text{Var}(X) = \lambda \text{ التباين}$$

التوزيع الإحصائي :-

و هو الشكل الذي تأخذه مجموعة البيانات، وشكل البيانات مهم جدا في تحليلها ووصفها وكخطوة تسبق قرار استخدام أي أسلوب احصائي .

ويرتبط التوزيع الاحصائي عادة بنوعين من البيانات المتصلة والمنفصلة، ويناسب النوع المنفصل المقاييس الاسمية والرتبية ، وهناك بعض المقاييس المنفصلة ثنائية أي انه لا يوجد بها الا قيمتين، وهي لا تسمى توزيعات طبيعية وانما تسمى توزيعات ثنائية ، ومن أهم مقاييس التوزيعات المنفصلة مقياس ذو الحدين وذلك عائد لان الاجابة على المقياس الاسمي اما نعم أو لا ، ولذلك غالبا ما يرمز لها في الحاسب بصفر (غياب الصفة) [ذكور - لا] أو ١ (وجود الصفة) [اناث - نعم] .

أما التوزيعات الاحصائية المتصلة فهي ذات أهمية كبيرة في العلوم الإحصائية وذلك لأن اغلب الاختبارات الاحصائية تتعامل مع هذا النوع من البيانات.

التوزيعات الاحتمالية للمتغيرات المتصلة:

هناك بعض التوزيعات الاحتمالية المتصلة لها دوال كثافة احتمال محددة ومنها:

- التوزيع الطبيعي
- التوزيع الطبيعي (القياسي) المعياري
- توزيع t

وسنقوم في هذه المحاضرة بتناول هذه التوزيعات بشيء من التوضيح والتفصيل:

وكما أوضحنا أن المتغير العشوائي المتصل x هو ذلك المتغير الذي يمكن أن يأخذ عدداً لا نهائياً من القيم المعلومة، واحتمال أن تقع x داخل أي فترة يمثلها مساحة التوزيع الاحتمالي (ويسمى أيضاً دالة الكثافة) داخل هذه الفترة، والمساحة الكلية تحت المنحنى (الاحتمال) تساوي 1

التوزيع الطبيعي

أ- التوزيع الطبيعي:

هو أفضل وأكثر التوزيعات الاحتمالية المتصلة استخداماً في النواحي التطبيقية، ومنها الاستدلال الإحصائي شاملاً التقدير، واختبارات الفروض، كما أن معظم التوزيعات يمكن تقريبها إلى هذا التوزيع .

والتوزيع الطبيعي هو توزيع احتمالي متصل، وهو جرسى الشكل ومتماثل حول الوسط الحسابي، ويمتد إلى مالا نهاية في الاتجاهين، ولكن معظم المساحة (الاحتمال) تتركز حول الوسط الحسابي .

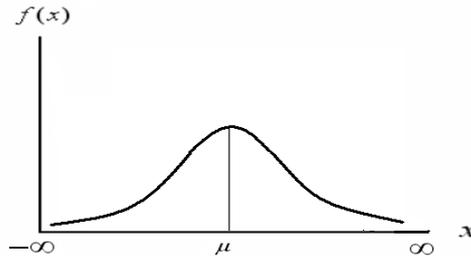
خصائص التوزيع الطبيعي:

يعتبر التوزيع الطبيعي من أهم أنواع التوزيعات الاحصائية المتصلة ومن خصائصه انه:

- توزيع جرسى أي يشبه الجرس.
- توزيع متصل
- توزيع متماثل حول الوسط
- الالتواء (الاطراف) والتفلطح (القمة) يساوي صفر.

- يحوي منوال ووسط ووسيط واحد وذات قيم متساوية بمعنى أن الجزء الذي على يمين الوسط مطابق للجزء الأيسر
- الذيلين الأيمن والأيسر يقتربان من الخط الأفقي ولكن لا تلامسه
- المساحة الكلية تحت المنحنى تساوي واحد صحيح
- منحنى دالة الاحتمال للتوزيع الطبيعي له خاصية شكل الجرس. ويتحدد شكل الجرس تماماً لأي توزيع طبيعي خاصة إذا علمنا الوسط الحسابي μ والانحراف المعياري σ لهذا التوزيع.
- تدل قيمة μ على مكان مركز الجرس، كما تدل σ على كيفية الانتشار.
- القيمة الصغيرة لـ σ تعني أن لدينا جرس طويل مدبب، والقيمة الكبيرة لـ σ تعني أن الجرس قصير ومفطح.

والشكل التالي يوضح ذلك:



والتوزيع الطبيعي وتطبيقاته الإحصائية ليس موضوعاً جديداً بل عرف منذ القرن السابع عشر الميلادي ومن أبرز الدراسات المعروفة تلك الدراسة البريطانية التي أخذت أطوال ٨٥٨٥ من الأفراد البريطانيين في القرن التاسع عشر وعمل هذا المنحنى وبالتالي تم اعتبار هذه العينة تمثل التوزيع الطبيعي.

معالم هذا التوزيع:

توجد معلمتين لهذا التوزيع هما :

$$\text{الوسط الحسابي : } E(x) = \mu \quad \text{والتباين : } \text{var}(x) = \sigma^2$$

ومن ثم يعبر عن توزيع المتغير بالرموز : $x \sim N(\mu, \sigma^2)$ ويعني ذلك أن المتغير العشوائي x يتبع التوزيع الطبيعي بمتوسط μ ، وتباين σ^2 .

التوزيع الطبيعي القياسي (المعياري) :-

- احتمال وقوع أية مشاهدة على بعد انحراف معياري واحد من الوسط الحسابي هو 0.6827
- احتمال وقوع أي مفردة على بعد إنحرافين معياريين من الوسط الحسابي هو 0.9545
- احتمال وقوع أية مفردة على بعد ثلاثة انحرافات معيارية من الوسط الحسابي هو 0.9973

استخدامات التوزيع الطبيعي القياسي:-

يستخدم التوزيع الطبيعي القياسي في التعامل مع الكثير من المشاكل العملية وإيجاد القيم الاحتمالية لها وإليك بعض الأمثلة على ذلك:

مثال:

افترض أن إدارة المرور بالأحساء وضعت جهازاً للرادار على طريق الدمام عند مدخل المبرز وذلك لضبط السيارات المسرعة في فترة معينة من اليوم، افترض أن X تمثل السرعة في الساعة للسيارات التي تمر بمدخل المبرز في فترة عمل الرادار، إذا كانت X تتوزع توزيعاً معتدلاً وسطه الحسابي 60 ميلاً وتباينه 25 ميلاً، أوجد التالي:

- نسبة السيارات التي تقل سرعتها عن 50 ميلاً في الساعة .
- نسبة السيارات التي تزيد سرعتها عن 65 ميلاً في الساعة.

- نسبة السيارات التي تكون سرعتها بين 60 ميلا و 70 ميلا في الساعة .
- عدد السيارات التي تكون سرعتها بين 60 ميلا و 77.45 ميلا من بين 10000 سيارة .

المحاضرة الرابعة : توزيعات المعاينة Sampling Distribution's

الاستدلال الإحصائي :-

تهتم نظرية العينات 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 والعلوم، ويشمل الاستدلال الإحصائي اختبار الفرضيات والتقدير .

ولكى يكون التقدير (واختبار الفروض) سليماً، ينبغي أن يبنى على عينة ممثلة للمجتمع، ويمكن تحقيق ذلك بالمعاينة العشوائية، حيث يكون لكل مفردة في المجتمع فرصة متكافئة للدخول في العينة .

العينة العشوائية:-

وهناك عدة طرق لأخذ العينات من المجتمع لاستخدامها في الاستدلال الإحصائي ومن أشهر هذه الطرق هي العينة العشوائية وهي العينة التي تكون لكل مفردة من مفردات المجتمع نفس فرصة الاختيار في العينة.

فمثلاً نستعين بعينه مسحوبة من المجتمع لتقدير معالم هذا المجتمع مثل متوسطة أو تباينه أو غير ذلك. أو أعطاء عينه من المرضى بارتفاع الضغط، مثلاً دواء معين ثم قياس ضغطهم قبل وبعد تناولهم لهذا الدواء لمعرفة ما إذا كان هذا الدواء مفيد في خفض الضغط أم لا.

المجتمع Population

أي مجموعات من المفردات تشترك في صفة أو صفات وتكون موضوع دراسة أو بحث فإن هذه المجموعة يطلق عليها إحصائياً 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 معينة خلال فترة زمنية محددة... 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 (لانهايتي) إذا كان لا يمكن حصر عدد أفراده مثل النجوم والكواكب أو الكائنات الحية بمياه المحيطات والأنهار

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

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

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

بعض مزايا أسلوب المعاينة:-

يتميز أسلوب المعاينة عن أسلوب الحصر الشامل بمزايا عديده منها:

١. 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.

٢. 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.
٣. 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.
٤. 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 شحنه من المفرعات مثلاً لابد وأن يتم على أساس العينة وبالمثل تحليل دم المرضى يتم على أساس عينه.

أقسام العينات:-

تنقسم العينات عادة إلى قسمين رئيسيين وهما عينات عشوائية وعينات غير عشوائية، وفيما يلي تفصيل لكل قسم منها:

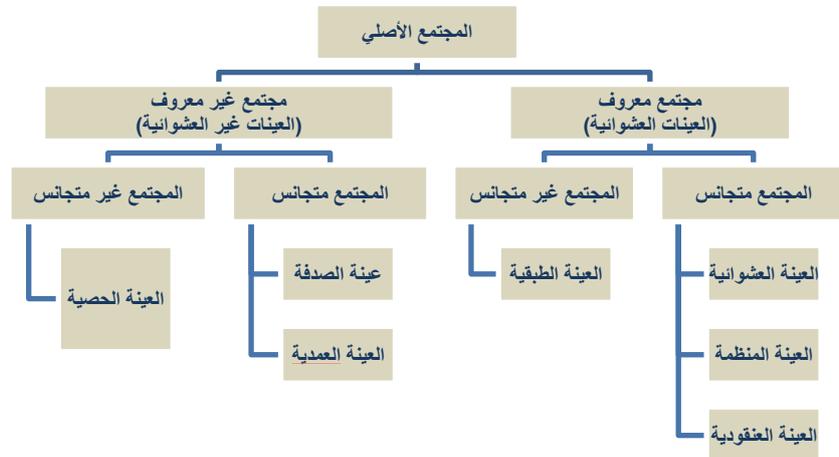
١. العينات العشوائية:

وهي تلك العينات التي يتم اختيار مفرداتها حسب خطه إحصائية لا يكون فيها للباحث أو لمفردات العينة دخل في اختيار أي مفرده 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 درجة عالية من دقة التمثيل للمجتمعات المسحوبة منها لذلك فهي الوسيلة الأساسية في حالة البحوث العلمية الدقيقة

٢. 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.

وسيتم فيما يلي استعراض لأهم أنواع العينات العشوائية والعينات غير العشوائية.



أ - العينات الاحتمالية:

العينة العشوائية	جميع عناصر المجتمع لها نفس الفرصة في الظهور في العينة
العينة الطبقيّة	يقسم المجتمع إلى طبقتين على الأقل ثم نختار العينة من كل منهما
العينة المنتظمة	نختار نقطة بداية من المجتمع ثم نختار العنصر الموجود على بعد ثابت من هذه النقطة
العينة العنقودية	يقسم المجتمع إلى مساحات أو أجزاء ثم نختار عشوائياً بعض هذه المساحات، ثم نختار جميع 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
العينة الحصية	يقسم المجتمع إلى أجزاء ثم نختار العينة من كل جزء من أجزاء المجتمع وفقاً للنسب المحددة

أخطاء البيانات الإحصائية: -

تتعرض البيانات الإحصائية التي يتم جمعها إلى نوعين من الأخطاء:

١. **خطأ التمييز أو التحيز:** وهو ذلك الخطأ الناتج عن مصادر متعددة، منها أخطاء في تصميم البحث أو التجربة أو أخطاء فنية أثناء جمع البيانات أو خلال العمليات الحسابية التي تتم على البيانات المتجمعة.. أخطاء التمييز تزداد بازدياد الفروق بين الإمكانات (المادية والفنية) اللازم توافرها لضمان أقصى درجة دقة ممكنة وبين الإمكانات الفعلية المتاحة للباحث.
٢. **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

وفيما يلي شرح لهذين الخطأين:

١- **خطأ التمييز أو التحيز:**

إذا سحبنا عدة عينات من مجتمع ما وحسبنا المتوسط الحسابي لكل عينة من هذه العينات ثم حسبنا المتوسط الحسابي لهذه المتوسطات فهذا المتوسط يجب أن يساوي المتوسط الحسابي الحقيقي للمجتمع المسحوبة منه هذه العينات، وفي حال وجود فرق بين المتوسطين فإن هذا الفرق يسمى **بخطأ التمييز أو التحيز**

أسباب خطأ التمييز أو التحيز:

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

٢- **خطأ المعاينة العشوائية Random Sampling Error**

عند اختيار العينة العشوائية هناك خطأ ينتج عن الاختلاف أو التشتت Variation بين قيم الوحدات التي تتكون منها العينة وتلك الوحدات التي لم تشأ الصدفة أن تدخلها في العينة وهذا الخطأ يسمى بخطأ المعاينة العشوائي

كيف نقلل من خطأ المعاينة العشوائي:

- زيادة حجم العينة
- طريقة الاختيار المناسب التي تقلل من اختلاف قيم الوحدات الإحصائية (كالأسلوب الطبقي أو العينة المنتظمة... الخ).

المعالم والإحصاءات:-

اعتاد البعض على معاملة القيم التي يحصل عليها من العينة وكأنها قيم مجتمعها، وهذا خطأ فادح. فلكي يستدل على خصائص مجتمع الدراسة تعتمد معادلات عديدة، ومتنوعة حسب نوع العينة.

فالمقاييس الإحصائية التي تحسب من بيانات مجتمع الدراسة بأكمله يطلق عليها معالم المجتمع (Parameters of population)

أما المقاييس الإحصائية التي تحسب من بيانات عينه 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 للانحراف المعياري للمجتمع بالرمز σ بينما يرمز للانحراف المعياري للعينة بالرمز S وهكذا.

الخطأ المعياري:

بمعرفة قيمة الانحراف المعياري لقيم العينة يمكن تقدير قيمة الخطأ المعياري في الانحراف المعياري للعينة باعتماد المعادلة الآتية:-

$$SE = \left[\frac{S}{\sqrt{n}} \right] \left[\sqrt{1 - \left(\frac{n}{N} \right)} \right] \quad \text{حيث ان:}$$

(N) = حجم مجتمع العينة و (n) = حجم العينة

وفق هذه المعادلة تؤخذ نسبة العينة إلى مجتمعها، وكلما كبرت هذه النسبة تحسن تمثيلها لمجتمع الدراسة، أما عندما يكون حجم مجتمع العينة مجهولاً، حينها تعتمد المعادلة الآتية:-

$$SE = \frac{S}{\sqrt{n}}$$

الخطأ المعياري = الانحراف المعياري لقيم العينة / جذر حجم العينة

أما عندما يكون حجم العينة أكثر من (١٠٠) فتعتمد المعادلة أدناه:-

$$SE = \frac{S}{\sqrt{n^2}}$$

الخطأ المعياري = الانحراف المعياري لقيم العينة / جذر (حجم العينة)^٢

إن الانحراف المعياري للتوزيع النظري لمتوسطات العينات يقيس خطأ المعاينة ويسمى بالخطأ المعياري للمتوسط، ومن الضروري التذكر دوماً أن متوسط المجتمع قيمة محددة تقع ضمن مجال محدد Certain Interval ، والباحث غير متأكد من قيمتها، ولكنه

يحسب احتمالية وجودها ضمن المجال المحدد وبمستوى ثقة إحصائية معلوم

مستوى الثقة وحدودها:

إذا أخذت جميع العينات المحتملة من مجتمعها فيتوقع أن تكون متوسطات العينات موزعة بالتساوي حول متوسط مجتمع الدراسة.

بعبارة أخرى، إن متوسط متوسطات العينات يساوي متوسط مجتمعها.

وتتوزع متوسطات العينات دائماً بصورة متمثلة Normal Distribution ، والذي يمتاز رياضياً بالابتعاد بنسب ثابتة عن المتوسط

مع كل درجة معيارية، وبالتالي تباينت متوسطات العينات المأخوذة منه فإنه يتوقع أن يقع متوسطه و باحتمالية قدرها كالتالي:

• مستوى ثقة إحصائية قدره (٦٨.٢٦%) أو باحتمالية قدرها (٠.٦٨٢٧) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات

العينات و (+ و -) درجة واحدة من الخطأ المعياري .

• مستوى ثقة إحصائية قدرها (٩٥%)، أو باحتمالية (٠.٩٥) يقع متوسط مجتمع الدراسة بين متوسط متوسطات العينات و (+ و -)

درجتان من الخطأ المعياري تقريباً.

• مستوى ثقة إحصائية قدرها (٩٩%) أو باحتمالية قدرها (٠.٩٩) يقع متوسط مجتمع الدراسة بين قيمة متوسط متوسطات العينات و

(+ و -) ثلاث درجات من قيمة الخطأ المعياري تقريباً.

• وتسمى هذه بمستويات الثقة Confidence Level و يعبر عنها بإشارة النسبة المئوية (%) بان تكون التقديرات صحيحة أو

باحتمالية (٠.٠١) أو (٠.٠٥) أن تكون خاطئة.

مثال :-

قام أحد الباحثين في مجال الزراعة بدراسة مائة مزرعة، فوجد أن متوسط مساحة المزرعة الواحدة (٥٣) هكتاراً، وانحراف

معيارى عن المتوسط بقيمة (٢٦) هكتاراً

أحسب حدود الثقة في تقدير متوسط مساحة المزرعة في منطقة الدراسة؟

حجم العينة:-

إن لحجم العينة أهمية كبيرة في تحديد الثقة بالنتائج، لذا من الضروري أن يسلط الضوء عليه بشيء من التفصيل وحسب التوزيعات

المعروفة للقيم، وسيتم هنا تناول نوعين من التوزيعات وهي:

١. التوزيع الطبيعي للقيم

٢. توزيع (ت) للقيم

(أ) التوزيع الطبيعي للقيم:

كلما كبر حجم العينة ازدادت دقة تمثيلها لمجتمعها واقترب توزيع القيم فيها من التوزيع الطبيعي (التمائل الجانبيين) وأصبحت عملية

الاستدلال أكثر دقة. وللتوضيح نورد مثلاً، إذا أريد معرفة نسبة طلبة قسم الإدارة إلى مجموع طلبة الكلية فإن عينة من عشرة طلبة

قد لا تفي بالغرض، ولكن عينة من مائة طالب تفي بالغرض حتماً. بعبارة أخرى، إن حجم العينة أساسي لإعطاء صورة عن مجتمع

الدراسة وليس النسبة المئوية للعينة قياساً بحجم مجتمعها. فكلما ازداد حجم العينة ازدادت الثقة بتقديرات خصائص المجتمع وصغرت

معه حدود الثقة.

(ب) توزيع (ت) للقيم :

من الضروري اخذ الحذر عندما يكون حجم العينة صغيرا اقل من (٣٠) وذلك لأنها تتطلب إجراءات خاصة عند التحليل. فعندما يكون حجم العينة أكثر من (٣٠) يتجه توزيع قيمها نحو التوزيع الطبيعي وبغض النظر عن التوزيع الحقيقي لقيم مجتمع الدراسة. وبالنسبة للعينات الصغيرة الحجم فان توزيع قيمها يتأثر طبيعياً بتوزيع قيم المجتمع المأخوذة منه. **وعندما يكون توزيع قيم المجتمع معروفاً أو متوقفاً أن لا يكون طبيعي حينها يجب اعتماد حجم كبير للعينة**. أما إذا كان توزيع قيم المجتمع طبيعياً عندها يمكن اخذ عينات بحجم صغير ويعتمد توزيع (ت) (T) في التحليل و المقارنة.

يتشابه توزيع قيم (ت) مع شكل الجرس بزيادة الحجم حتى يتطابق معه عندما يتعدى العدد (٣٠) فشكل توزيع (ت) للقيم لا يختلف كثيراً عن التوزيع الطبيعي إلا في الأعداد القليلة، وكلاهما متماثل النصفين لذا يعتمد كبدل له في القيم القليلة العدد. ولتوزيع (ت) جداول للقيم الحرجة منظمة على شكل اسطر اعتماداً على درجة الحرية التي تقاس بـ (حجم العينة - ١). أما الأعمدة فتمثل درجة الاحتمالية Probability، وتتناقص القيم الحرجة بتزايد درجة الحرية (حجم العينة). **ودرجة الحرية تفضل على حجم العينة في الأحجام الصغيرة للعينة لأنها تقلل من الانحياز في تقدير خصائص مجتمع الدراسة.**

العوامل المحددة لحجم العينة:

درجة التباين في خصائص مجتمع الدراسة: يلعب التباين في خصائص مجتمع الدراسة دوراً مهماً في تحديد درجة دقة نتائج العينة، فكلما كان التباين كبيراً تطلب الأمر زيادة حجم العينة ليكون تمثيلها للتباين في المجتمع صحيحاً.
طريقة التحليل المعتمدة: عند إقرار حجم العينة، من الضروري تحديد الحجم الأصغر المقبول للعينة في المجاميع الثانوية ضمن مجتمع الدراسة، إذ أن بعض الاختبارات الإحصائية تتطلب عدداً معيناً كحد أدنى لكل فئة أو صنف لتكون النتائج ذات معنى.

حجم المعلومات المطلوبة: فكلما كانت المعلومات المطلوبة من العينة (الواحدة) كثيرة وتفصيلية كان حجم العينة صغيراً، ما لم يكن المشروع البحثي كبيراً وتتوفر له المصادر البشرية والمادية اللازمة. **إن الدقة في المعلومات المطلوبة من العينة أهم بكثير من حجم العينة**، فحجم العينة لا يحدد حجم مجتمع الدراسة فقط، بل وبالذقة المتوخاة والتفاصيل المطلوبة
المصادر المالية والبشرية المتوفرة: تتطلب الدراسة الميدانية توفر مصادر مالية وبشرية لتغطية تكاليفها التي تكون في الغالب باهظة لتأثيراتها على تحديد حجم منطقة الدراسة، مجتمع الدراسة وبالتالي حجم العينة. إن مضاعفة حجم العينة يتطلب زيادة في كمية المصادر المالية والجهد البشري

حدود الثقة في تقديرات خصائص مجتمع الدراسة: لزيادة الدقة في النتائج يعتمد البعض إلى تقليص حدود الثقة (المدى الذي يفترض أن يقع ضمنه المعدل المتوقع للمجتمع). إن إنقاص حدود الثقة من (٦%) إلى (٤%) يتطلب زيادة حجم العينة بنسبة (٢٢٥%)، وكلما كان المدى كبيراً كان حجم العينة صغيراً، والعكس صحيح.
حالات الإخفاق وعدم الاستجابة: العامل الآخر الذي يحدد حجم العينة هي حالات الإخفاق في الحصول على المعلومات وعدم الاستجابة أو المعلومات غير الوافية

المحاضرة السادسة**التقدير الإحصائي**

التقدير هو عملية استنتاج أو تقدير أحد معالم المجتمع (مثل الوسط الحسابي أو الانحراف المعياري) بناءً على بيانات عينة مسحوبة من المجتمع .

وهناك نوعان (أو أسلوبان) للتقدير:

- الأول تقدير النقطة (أو القيمة الواحدة)
- الثاني تقدير الفترة (أو فترة التقدير أو الثقة)

١- التقدير بنقطة :-

التقدير بنقطة يعني أن نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة. فمثلاً لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدخل نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة. وكمثال آخر لو أخذنا نسبة الناخبين في العينة الذين يؤيدون مرشحاً معيناً كتقدير لهذه النسبة في المجتمع نكون حصلنا على تقدير نقطة للنسبة في مجتمع الناخبين .

٢- التقدير بفترة :-

أما التقدير بفترة فنحصل من خلاله على مدى Range أو فترة تتحدد بحددين (حد أدنى وحد أعلى) - نحصل عليهما من العينة. ونلاحظ هنا أن فترة التقدير (أو تقدير الفترة) تحتوي على أكثر من قيمة بل قد يكون عدد القيم غير محدود أو لا نهائياً في كثير من الحالات.

مثلاً:

إذا قدرنا أن الوسط الحسابي لأعمار الناخبين يتراوح بين (6 - 40) و (6 + 40) سنة أي يتراوح بين 34 سنة كحد أدنى و 46 سنة كحد أعلى نكون قد حصلنا على تقدير فترة للوسط الحسابي لأعمار الناخبين في المجتمع - ونلاحظ أن هذه الفترة (34, 46) تحتوي على عدد لا نهائي من الأعمار، بمعنى أن العدد لا يقتصر فقط على الأعداد الصحيحة والتي تشمل السنوات، ولكنها تشمل أيضاً كسور السنوات، والأيام والشهور، والساعات.. الخ

تقدير الوسط الحسابي للمجتمع :-

أ - احسب الوسط الحسابي للعينة

ب- احسب الخطأ المعياري للوسط والذي يساوي: $\frac{\sigma}{\sqrt{n}}$

ج - أضرب الخطأ المعياري للوسط في معامل الثقة المناسب (أو الدرجة المعيارية) حسب درجة الثقة المطلوبة أي أحسب: $Z \frac{\sigma}{\sqrt{n}}$

د- فعندما نطرح حاصل الضرب السابق من الوسط الحسابي للعينة نحصل بالتالي على الحد الأدنى لفترة التقدير، وعندما نجمع حاصل الضرب مرة أخرى على الوسط الحسابي للعينة نحصل بالتالي على الحد الأعلى لفترة التقدير.

$$\mu = \bar{x} \pm z \frac{\sigma}{\sqrt{n}}$$

مع ملاحظة أنه إذا كان الانحراف المعياري للمجتمع غير معروف - وهو غالباً ما يحدث في الواقع - فيمكن استخدام الانحراف المعياري للعينة S بدلاً منه طالما كان حجم العينة كبيراً بدرجة كافية وتصبح فترة تقدير الوسط الحسابي للمجتمع كما يلي :

$$\hat{\mu} = \bar{X} \pm Z \frac{S}{\sqrt{n}}$$

ولإيضاح هذه النقطة بشيء من التفصيل نأخذ المثال التالي :

معامل الثقة Z	درجة الثقة
1	68.26%
1.65	90%
1.96	95 %
2	95.44%
2.58	99%

تحديد حجم العينة لتقدير الوسط الحسابي للمجتمع :-

يعتبر تحديد حجم العينة المناسب من المشاكل المهمة والشائعة التي تواجه الباحثين في مختلف المجالات، وبالذات عند دراسة الظواهر السياسية والاجتماعية... الخ، ويختلف تحديد حجم العينة باختلاف الهدف من التقدير.

فإذا كان المطلوب هو تقدير الوسط الحسابي للمجتمع، فإن فترة تقدير الوسط هي كما سبق وأن أوضحنا : $\hat{\mu} = \bar{X} \pm Z \frac{\sigma}{\sqrt{n}}$

ومنها نجد أن حجم العينة يأخذ الشكل التالي : $n = \frac{Z^2 \sigma^2}{e^2}$

Z = هو معامل الثقة (أو الدرجة المعيارية) المقابل لدرجة الثقة المطلوبة، ونحصل عليها من جدول التوزيع الطبيعي المعياري.
= هو تباين المجتمع (أو هو مربع الانحراف المعياري).

e = هو أقصى خطأ مسموح به في تقدير الوسط، وهو عادة ما يحدده الباحث، وتتوقف قيمته على أهمية الموضوع أو الظاهرة السياسية المراد دراستها، ومدى الدقة المطلوبة في التقدير، ويسمى اختصاراً "الخطأ في تقدير الوسط".

ولتوضيح كيفية تحديد حجم العينة المناسب عند تقدير الوسط الحسابي للمجتمع، نأخذ المثال التالي :

مما سبق :

في حالة تقدير النقطة نحصل على قيمة واحدة من العينة، وتستخدم هذه القيمة الواحدة كتقريب أو كتقدير لمعلمة المجتمع المجهولة.

فمثلاً لو أخذنا الوسط الحسابي للدخل في العينة كتقدير لمتوسط الدخل نكون قد حصلنا على تقدير نقطة لمتوسط دخل الدولة.

أما في حالة تقدير الفترة أو فترة التقدير فنحصل على مدى Range أو فترة تتحدد بحددين (حد أدنى وحد أعلى) - نحصل عليهما من العينة. ونلاحظ هنا أن فترة التقدير (أو تقدير الفترة) تحتوي على أكثر من قيمة بل قد يكون عدد القيم غير محدود أو لا نهائياً في كثير من الحالات.

فترات الثقة للمتوسط باستخدام توزيع t :-

تناولنا فيما سبق التقدير الإحصائي للوسط الحسابي للمجتمع في الحالات التي يكون فيها الانحراف المعياري للمجتمع معلوماً، و (أو) أن العينة كبيرة بدرجة كافية. ولكن إذا كانت العينة صغيرة بمعنى أن حجمها أقل من (30) مفردة، والانحراف المعياري للمجتمع الطبيعي غير معلوم، فإن التوزيع الإحصائي المتبع في مثل هذه الحالات هو ما يطلق عليه "توزيع t"

ولعل الاختلاف الأساسي بين توزيع t والتوزيع الطبيعي هو أن الانحراف المعياري للعينة هو المستخدم في الأول بدلاً من الانحراف المعياري للمجتمع في الثاني، وفيما عدا ذلك فالتوزيعان متماثلان وكلما زادت قيمة n كلما اقترب توزيع t من توزيع z ويعتمد توزيع t على ما يعرف بدرجات الحرية DEGREES OF FREEDOM .

درجات الحرية :

تعرف درجات الحرية بأنها عدد المشاهدات المستقلة في العينة والتي تساوي حجم العينة مطروحاً منه عدد القيود أو معالم المجتمع التي يتم تقديرها من بيانات العينة. وكمثال مبسط لشرح فكرة درجات الحرية نفترض أن لدينا 3 قيم واشترطنا أن مجموع القيم يساوي 10 فإن لدى الباحث في هذه الحالة حرية في اختيار الرقم الأول (وليكن 2) والثاني (وليكن 3) لذلك فإن قيمة الثالثة لا بد وأن تكون (5) بالتالي نستطيع القول بأن درجة الحرية المتاحة لدى الباحث هي (2) أي $2 = 3 - 1$ أي أن درجات الحرية في هذه الحالة هي : $n - 1$

حيث n تساوي حجم العينة (والتي تساوي في المثال السابق 3) والرقم (1) والذي طرحناه يعني الشرط الذي يحتم أن مجموع القيم = 10 وبصفة عامة إذا كان عدد القيود k فإن درجات الحرية تساوي $n - k$

وفترة الثقة 95% لوسط المجتمع غير المعلوم عند استخدام توزيع t هي:

$$P\left(\bar{X} - t \frac{s}{\sqrt{n}} < \mu < \bar{X} + t \frac{s}{\sqrt{n}}\right) = 0.95 \quad (4.7)$$

حيث تشير t الى قيمة t التي تقع عندها 2.5% من المساحة الكلية للمنحنى عند كل طرف (عند درجات الحرية المستخدمة)، وتستخدم s/\sqrt{n} بدلاً من σ/\sqrt{n}

شروط توزيع t :

ويمكن تحديد الشروط الثلاثة لاستخدام توزيع t كما يلي:

1. أن يكون المجتمع المسحوبة منه العينة له توزيع طبيعي.
2. والانحراف المعياري للمجتمع غير معروف (أو مجهول).
3. والعينة صغيرة (حجمها أقل من 30 مفردة).

تقدير فترة النسبة للمجتمع

(فترة الثقة للنسبة)

إن تقدير النسبة في المجتمع تعتبر من الحالات المهمة لقياس الظواهر السياسية، وبالذات الوصفية منها كقياس اتجاهات الرأي العام، وقياس نسبة قتلى الحروب، ونسبة الدول التي أوفت بالتزاماتها في المنظمات الدولية أو الإقليمية... وغيرها ونظراً لأنه من الصعوبة بمكان في كثير من الأحيان حساب هذه النسبة مباشرة من المجتمع، فإننا غالباً ما نلجأ لتقدير هذه النسبة من عينة عشوائية مسحوبة من هذا المجتمع.

خطوات تقدير النسبة في المجتمع:

فلو افترضنا أن نسبة المؤيدين للسياسة الاقتصادية التي تنتهجها دولة ما هي P وأن العينة العشوائية كبيرة بدرجة كافية وأن نسبة مؤيدي هذه السياسة في العينة هي \hat{P} فتقدير النسبة في المجتمع تكون كما يلي:

$$P = \hat{P} \pm Z \sqrt{\frac{\hat{P}(1 - \hat{P})}{n}}$$