دولة ليبيا
وزارة التعليم
مركز المناهج التعليمية والبحوث التربوية

أسس الإحصاء

للسّنة الثّالِثةَ بمرحلة التّعليم الثانوي

(القسم العلمي)

تأليف

أ. ناجي رشيد الكيخيا

مراجعة علمية

د. أحمد محمد مامي

مراجعة لغوية

أ. محمد الشلطامي

1440 – 1441 هـ

2019 – 2020 م

المقدمة:

ينقسم علم الإحصاء إلى فرعين رئيسين هما، الإحصاء الوصفي والإحصاء الاستدلالي. وقد تم التطرق للإحصاء الوصفي في كتاب الإحصاء للسنة الثانية وسنتعرض في هذا الكتاب لموضوع الإحصاء الاستدلالي، وهو يشمل الطرق والأساليب الإحصائية الخاصة بكيفية استخلاص النتائج واتخاذ القرارات المناسبة التي تعمم على الكل وهو ما يعرف بالمجتمع، بدراسة البيانات التي نحصل عليها من جزء من هذا الكل، وهو ما يعرف بالعينة. وينقسم الإحصاء الاستدلالي إلى قسمين هما، التقدير واختبارات الفروض، ويعتمد الإحصاء الاستدلالي اعتماداً كبيراً على ما يسمى بنظرية الاحتمالات، فباستخدامها يستطيع الباحث أو متخذ القرار تحديد احتمال الخطأ الممكن الوقوع فيه نتيجة دراسة الجزء بدلاً من الكل.

يحتوي هذا الكتاب على (6) فصول، نعرض في الفصل الأول المفاهيم الأساسية لنظرية الاحتمالات.

أما الفصل الثاني فيتناول تعريف المتغير العشوائي بنوعيه وهما المتغير العشوائي المنفصل والمتغير العشوائي المتصل، ونناقش التوزيع الاحتمالي المنفصل والتوزيع الاحتمالي المتصل، ونتعرض لكيفية حساب بعض المقاييس الإحصائية الهامة التي تُستخدم لوصف التوزيعات الاحتمالية.

في الفصل الثالث، نعرض بعض التوزيعات الاحتمالية المنفصلة والمتصلة الهامة والأكثر استعمالاً في التطبيقات الإحصائية، فمن التوزيعات الاحتمالية المنفصلة سنتعرض لتوزيع ذات الحدين، وتوزيع بواسون. وسنتناول بإسهاب أهم توزيع احتمالي متصل وهو التوزيع الطبيعي ثم سنتناول توزيع (**t**).

وسنعرض في الفصل الرابع توزيعات المعاينة أي التوزيعات الاحتمالية الخاصة بإحصاءات، ونتناول فيه بإسهاب توزيع المعاينة للوسط الحسابي للعينة.

أما الفصل الخامس فيتناول التقدير الإحصائي بنوعيه وهما التقدير بقيمة والتقدير بفترة، ونتعرض لفترة الثقة للوسط الحسابي للمجتمع.

ويتناول الفصل السادس اختبارات الفروض، وفيه نشرح المقصود بالفروض الإحصائية وأنواع الأخطاء التي يتعرض لها متخذ القرار، ثم نناقش اختبارات خاصة ببعض المعالم الهامة.

أرجو أن أكون قد وفقت في تقديم مادة هذا الكتاب، بأكبر قدر ممكن من التبسيط.

والله الموفق

الفصل الأول
نظرية الاحتمالات

نظرية الاحتمالات هي أساس الإحصاء الاستدلالي، فتمدنا بالطرق والأساليب الرياضية التي تساعدنا للوصول إلى أفضل الاستنتاجات والقرارات التي تخضع لدرجة من عدم التأكد، وذلك بسبب دراسة الجزء للاستدلال على صفات الكل. وقبل التطرق لنظرية الاحتمالات سنعرض بعض المصطلحات الإحصائية الهامة التي لها علاقة مباشرة بموضوع الاحتمالات، وهذه المصطلحات هي:

(1−1) التجربة العشوائية:

تُعرف التجربة العشوائية بأنها أية عملية قد تعطى نتائج مختلفة حتى إذا أعيدت تحت نفس الظروف، ولا يمكن أن نتنبأ أو نحدد بشكل أكيد نتيجتها قبل إجرائها، ولكننا نعرف مسبقـًا كل النتائج التي يمكن الحصول عليها.

مثال (1−1)

عند إلقاء قطعة نقدية في الهواء وتركها تعود، في هذه العملية نستطيع مسبقـًا معرفة كل النتائج الممكن الحصول عليها وهي وجه أو ظهر، ولكن لا نعرف مسبقـًا أي نتيجة من هذه النتائج ستظهر حتمـًا، فهذه العملية يطلق عليها تجربة عشوائية.

مثال (1−2)

عند إلقاء مكعب نرد في الهواء وتركه يعود، في هذه العملية نستطيع مسبقـًا معرفة كل النتائج الممكن الحصول عليها وهي الإعداد التالية **6,5,4,3,2,1**، ولكن لا نستطيع مسبقـًا تحديد أي نتيجة من هذه النتائج سنحصل عليها، فهذه العملية يطلق عليها تجربة عشوائية.

مثال (1-3)

إذا كان لدينا **10** طلبة، وأردنا أن نختار منهم طالباً واحداً عشوائياً، حيث المقصود بالاختيار العشوائي هو اختيار الطالب بطريقة تضمن إعطاء نفس الفرصة لكل طالب من الطلبة العشرة ليكون هو الطالب المختار، أي يجب أن يكون الاختيار خاضعاً لعامل الصدفة المطلقة دون تدخل العامل البشري فيه، ويتم ذلك بإعطاء رقم لكل طالب، وتكتب هذه الأرقام على بطاقات متماثلة تماماً، ثم نضع كل البطاقات في وعاء ونخلطها جيداً ثم نسحب ونحن مغمضي العينين بطاقة، فالرقم الذي على البطاقة هو رقم الطالب المختار، في هذه العملية نعرف مسبقاً انه سيظهر رقم أحد الطلبة العشرة، ولكننا لا نستطيع أن نحدد مسبقاً رقم أي طالب من هؤلاء الطلبة سيظهر، وبالتالي فهذه العملية تسمى تجربة عشوائية .

وتعتمد نظرية الاحتمالات على التجارب العشوائية، وبالتالي ستكون التجارب التي نتعامل معها في موضوع هذا الكتاب كلها تجارب عشوائية.

(1-2) فراغ العينة :

فراغ العينة لتجربة هو المجموعة التي تشمل كل النتائج التي يمكن الحصول عليها من إجراء هذه التجربة .

عند القيام بأي تجربة، تظهر لنا نتيجة واحدة فقط من النتائج التي يشملها فراغ العينة لهذه التجربة، فلا نستطيع الحصول على أكثر من نتيجة من هذه النتائج في نفس الوقت، وبالطبع يختلف فراغ العينة من تجربة لأخرى.

وعادة يرمز للمجموعة التي تمثل فراغ العينة بالحرف **S**، وهي تقابل الفئة الشاملة في موضوع المجموعات، وتسمى كل نتيجة من النتائج التي يشملها فراغ العينة عنصر أو نقطة فراغ العينة.

مثال (1-4):

في تجربة إلقاء قطعة نقدية، سنحصل على وجه أو ظهر، فإذا رمزنا للوجه بالحرف **H** ورمزنا للظهر بالرمز **T**، ففراغ العينة لهذه التجربة سنعبر عنه بالمجموعة التالية:

$$S = \{ H , T\}$$

مثال (1-5):

في تجربة إلقاء مكعب نرد مرة واحدة، كل النتائج التي يمكن الحصول عليها هي الأعداد: **6,5,4,3,2,1** وبذلك فإن فراغ العينة لهذه التجربة هو:

$$S= \{1,2,3,4,5,6\}$$

مثال (1-6):

في تجربة إلقاء قطعة نقود مرتين، فكل النتائج التي يمكن أن نحصل عليها هي وجه في الرمية الأولى ووجه في الرمية الثانية **(H H)** ، أو وجه في الرمية الأولى وظهر في الرمية الثانية **(HT)**، أو ظهر في الرمية الأولى ووجه في الرمية الثانية **(T H)** أو ظهر في الرمية الأولى وظهر في الرمية الثانية **(T T)**. النتيجة المكتوبة ناحية اليسار هي نتيجة الرمية الأولى والنتيجة المكتوبة ناحية اليمين هي نتيجة الرمية الثانية، إذن فراغ العينة لهذه التجربة:

$$S= \{ HH , HT ,TH ,TT \}$$

ملاحظة:

فراغ العينة لتجربة إلقاء قطعتي نقود هو نفسه فراغ العينة لتجربة إلقاء قطعة نقود مرتين، فنتيجة القطعة الأولى ستكون مقابلة لنتيجة الرمية الأولى، ونتيجة القطعة الثانية ستكون مقابلة لنتيجة الرمية الثانية.

مثال (1-7):

إذا اخترنا عشوائيا، ثلاث وحدات منتجة من آلة معينة، لفحصها ما إذا كانت تالفة أو غير تالفة، فإذا رمزنا للوحدة التالفة بالحرف **D**، وللوحدة غير التالفة بالحرف **N**، فسنعبر عن النتائج كما يلي، فإذا كانت الوحدات الثلاثة غير تالفة فسنكتب نتيجة الفحص **N NN**، وإذا كانت الوحدة الأولى تالفة والثانية والثالثة غير تالفتين فسنكتب النتيجة **DNN**، وهكذا .....، وبالتالي سيكون فراغ العينة لهذه التجربة وهي تجربة فحص ثلاث وحدات منتجة كما يلي :

$$S=\{\ NNN\ ,\ DNN\ ,\ NDN\ ,\ NND\ ,\ DDN\ ,\ DND\ ,\ NDD\ ,\ DDD\ \}$$

مثال (1-8):

في تجربة إلقاء مكعبي نرد، فراغ العينة لهذه التجربة سيكون كما يلي :

$$S=\left\{\begin{matrix} (1,1) & (1,2) & (1,3) & (1,4) & (1,5) & (1,6) \\ (2,1) & (2,2) & (2,3) & (2,4) & (2,5) & (2,6) \\ (3,1) & (3,2) & (3,3) & (3,4) & (3,5) & (3,6) \\ (4,1) & (4,2) & (4,3) & (4,4) & (4,5) & (4,6) \\ (5,1) & (5,2) & (5,3) & (5,4) & (5,5) & (5,6) \\ (6,1) & (6,2) & (6,3) & (6,4) & (6,5) & (6,6) \end{matrix}\right\}$$

حيث العدد المكتوب ناحية اليسار في كل نتيجة هو العدد الذي يظهر على المكعب الأول والعدد المكتوب ناحية اليمين هو العدد الذي يظهر على المكعب الثاني، فمثلاُ النتيجة **(2,3)** تعني ظهور العدد **2** على المكعب الأول وظهور العدد **3** على المكعب الثاني وتقرأ هذه النتيجة (اثنان، ثلاثة).

(1−3) الحدث:

في أية تجربة عشوائية قد نكون راغبين في ظهور نتائج معينة من مجمل النتائج التي يمكن الحصول عليها من هذه التجربة، أي من النتائج التي يشملها فراغ العينة لهذه التجربة، دون النتائج الأخرى، وهذه النتائج المرغوب ظهورها أي حدوثها يطلق عليها مصطلح حدث، ويعبر عن أي حدث بمجموعة، ويرمز عادة للمجموعة التي تمثل الحدث بأحد الحروف **A , B , C,** ... مع عدم استعمال الحرف **S** لأنه يستعمل كرمز لفراغ العينة.

وبما أن النتائج التي تشملها المجموعة الممثلة لأي حدث هي جزء من النتائج الكلية التي يمكن الحصول عليها من التجربة، فبالتالي المجموعة التي تمثل أي حدث يجب أن تكون مجموعة جزئية من فراغ العينة. ومن هنا يعرف الحدث كما يلي:

الحدث هو مجموعة جزئية من فراغ العينة S .

4 – أنواع الأحداث:

1 – الحدث البسيط :

عندما تحتوي المجموعة الجزئية التي تمثل الحدث على نتيجة واحدة فقط من نتائج فراغ العينة، يكون الحدث بسيطاً.

مثال (1-9):

إذا القينا مكعب نرد، ونرغب في ظهور العدد **5**، ففراغ العينة:

$$S= \{1,2,3,4,5,6\}$$

والحدث هنا هو ظهور العدد **5**، وإذا رمزنا له بالحرف A، فنعبر عنه كما يلي:

$$A= \{ 5 \}$$

ونلاحظ أن المجموعة الجزئية التي تمثل هذا الحدث تحتوي على نتيجة واحدة فقط من نتائج فراغ العينة، إذن فهذا الحدث هو حدث بسيط.

1- الحدث المركب:

عندما تحتوي المجموعة الجزئية التي تمثل الحدث على أكثر من نتيجة من نتائج فراغ العينة، يكون الحدث مركباً.

مثال (1-10):

إذا القينا مكعب نرد، ونرغب في ظهور عدد أكبر من 3، ففراغ العينة:

$$S= \{1,2,3,4,5,6\}$$

والحدث هنا هو ظهور عدد أكبر من 3، وإذا رمزنا له بالحرف A، فسنعبر عنه كما يلي:

$$A= \{ 4 , 5 , 6 \}$$

ونلاحظ أن المجموعة الجزئية التي تمثل هذا الحدث تحتوي على أكثر من نتيجة من نتائج فراغ العينة، إذن فهذا الحدث هو حدث مركب.

2- الحدث المؤكد:

يسمى الحدث حدثًا مؤكداً عندما تحتوي المجموعة الجزئية التي تمثل الحدث على كل نتائج فراغ العينة ، أي أن المجموعة الجزئية التي تمثل الحدث المؤكد تساوى فراغ العينة، فإذا رمزنا للحدث المؤكد بالرمز **A** فإن: **S= A** ، ويعني ذلك أن كل نتائج التجربة تحقق الحدث المرغوب فيه وبالتالي فمن المؤكد أن يتحقق، ومن هنا يطلق عليه الحدث المؤكد.

مثال (1-11):

إذا ألقينا مكعب نرد، ونرغب في ظهور عدد أقل من **7**، ففراغ العينة لهذه التجربة:

$$S= \{1,2,3,4,5,6\}$$

والحدث المطلوب هنا هو ظهور عدد أقل من **7**، فإذا رمزنا لهذا الحدث بالرمز **A**، فسنعبر عنه كما يلي:

$$A= \{1,2,3,4,5,6\}$$

ونلاحظ أن المجموعة الجزئية التي تمثل هذا الحدث تساوي فراغ العينة. أي أن كل النتائج الممكن الحصول عليها من هذه التجربة تحقق هذا الحدث، حيث أن من المؤكد أن نحصل على عدد أقل من **7** عند إلقاء مكعب نرد، لأن كل الأعداد الموجودة على مكعب النرد هي أعداد أقل من **7**، ولذلك يسمى هذا الحدث بالحدث المؤكد.

3- الحدث المستحيل:

يسمى الحدث حدثًا مستحيلاً عندما تكون المجموعة الجزئية التي تمثل الحدث خالية من العناصر، أي لا توجد أية نتيجة من نتائج فراغ العينة تحقق الحدث المطلوب، أي أن كل النتائج الممكن أن نحصل عليها من التجربة لا تحقق الحدث المرغوب فيه. وبالتالي فمن المستحيل أن يتحقق هذا الحدث عند إجراء التجربة، ولذلك سمى بالحدث المستحيل. ويرمز للحدث المستحيل بالمجموعة الخالية $\emptyset$، وبالطبع هذا لا يخل بتعريف الحدث، فالمجموعة الخالية هي مجموعة جزئية لأية مجموعة .

مثال (1-12):

إذا ألقينا مكعب نرد، ونرغب في الحصول على العدد **9**، ففراغ العينة لهذه التجربة:

$$S= \{1,2,3,4,5,6\}$$

والحدث هنا هو الحصول على العدد **9**، فإذا رمزنا لهذا الحدث بالرمز **A**، فسنعبر عنه كما يلي:

$$A = \{\quad\} = \emptyset$$

ونلاحظ أن المجموعة الجزئية التي تمثل هذا الحدث مجموعة خالية من العناصر. وذلك لأن كل النتائج الممكن الحصول عليها من هذه التجربة لا تحقق هذا الحدث، فمن المستحيل أن نحصل على العدد 9 عند إلقاء مكعب نرد، لأن العدد **9** ليس من الأعداد الموجودة على مكعب النرد، وبالتالي فهذا الحدث هو حدث مستحيل.

5- الحدث المكمل:

لكل حدث حدثاً مكملاً له، فالحدث المكمل للحدث **A** مثلاً، هو الحدث الذي يحتوي على نتائج فراغ العينة التي لا يحتويها الحدث **A** ، وذلك كما هو موضح في شكل **(1-1)** ويرمز لمكمل الحدث **A**، وبالرمز **A`**، ومن تعريف الحدث المكمل نستنتج أن أي حدث مكمل يجب أن يحقق الشرطين التاليين:

أ) اتحاد أي حدث والحدث المكمل له يجب أن يعطي فراغ العينة، أي أن:

$$A \cup A` = S$$

ب) تقاطع أي حدث والحدث المكمل له يجب أن يعطي مجموعة خالية، أي أن:

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

شكل (1-1)

مثال (1-13):

إذا القينا مكعب نرد مرة واحدة، وكان الحدث **A** هو الحصول على عدد زوجي، فما هو الحدث المكمل للحدث **A** ؟

الحل:

فراغ العينة لهذه التجربة: **S= {1,2,3,4,5,6}**

الحدث **A**: **A= { 2, 4, 6 }**

إذن الحدث المكمل للحدث A هو الحدث الذي يحتوي على كل عناصر فراغ العينة غير الموجودة في A وهي الأعداد الفردية 1 ، 3 ، 5 ، ونعبر عنه كما يلي :

$$A^{`}=\{\ 1\ ,3\ ,5\ \}$$

ونلاحظ أن :

$$A\cup A^{`}=\{1\ ,2\ ,3\ ,4\ ,5\ ,6\ \}=S$$
$$A\cap A^{`}=\{\quad\}=\emptyset$$

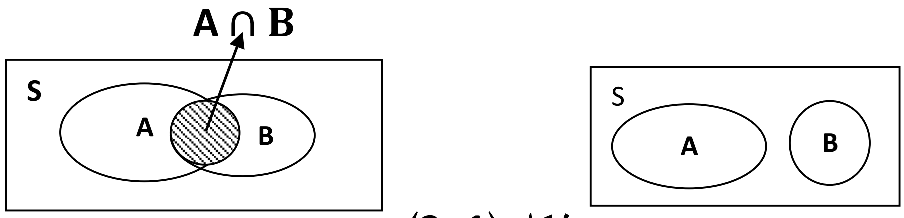
6- الأحداث المتنافية:

تكون الأحداث متنافية إذا كان ظهور أحدها يمنع الأحداث الأخرى من الظهور، أي لا يمكن أن نحصل عليها كلها في نفس الوقت. فمثلاً إذا كان الحدثان B , A حدثين متنافيين، فيعني ذلك أن ظهور الحدث A يمنع الحدث B من الظهور، وظهور الحدث B يمنع الحدث A من الظهور، أي لا نستطيع أن نحصل على الحدثين B,A معـاً في نفس الوقت. ونفهم من ذلك أنه لا توجد نتيجة من نتائج التجربة تحقق الحدثين معـاً في نفس الوقت، أي لا توجد عناصر مشتركة بين المجموعة الجزئية التي تمثل الحدث A والمجموعة الجزئية التي تمثل الحدث B، أي تقاطعهما مجموعة خالية:

$$A\cap B=\emptyset$$

ويكون الحدثان غير متنافيين، إذا كان من الممكن أن يظهراً معـاً في نفس الوقت، أي توجد نتائج مشتركة بينهما. أي أن تقاطعهما مجموعة غير خالية، وذلك كما هو موضح في شكل (1-2).

حدثان متنافيان **حدثان غير متنافيين**



شكل (1−2)

مثال)1-14(:

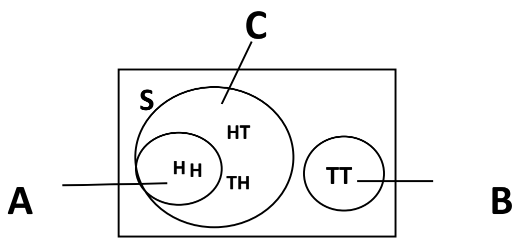
في تجربة إلقاء قطعتي نقود، إذا كان الحدث **A** هو الحصول على وجهين، والحدث **B** هو الحصول على ظهرين، والحدث **C** هو الحصول على وجه أو أكثر . فأي حدثين متنافيين وأيهما غير متنافيين؟

الحل:

فراغ العينة لهذه التجربة: **S = { HH , HT , TH , TT }**

والمجموعات الجزئية التي تمثل الأحداث هي:

A = { HH }
B = { TT }
C = {HH , HT , TH }



شكل (1−3)

بما أنه لا توجد نتيجة تحقق الحدثين **A , B** في نفس الوقت، أي تقاطعهما يعطي مجموعة خالية :

$$A \cap B = \emptyset$$

إذن الحدثان **B , A** حدثان متنافيان .

بما انه توجد نتيجة تحقق الحدثين **C , A** في نفس الوقت، أي تقاطعهما ليس مجموعة خالية :

$$A \cap C = \{HH\}$$

إذن الحدثان **C ,A** حدثان غير متنافيين .

بما انه لا توجد نتيجة تحقق الحدثين **C , B** في نفس الوقت، أي تقاطعهما مجموعة خالية :

$$B \cap C = \emptyset$$

إذن الحدثان **C , B** حدثان متنافيان . وشكل **(1-3)** يوضح ذلك.

7-الأحداث المستقلة:

يكون الحدثان **B , A** حدثين مستقلين إذا كان ظهور أحدهما لا يؤثر ولا يتأثر بظهور أو عدم ظهور الآخر . أي لا توجد أية علاقة بينهما.

فمثلاً عند إلقاء قطعة نقود مرتين، فظهور وجه في الرمية الأولى لا يؤثر ولا يتأثر بظهور وجه في الرمية الثانية، أي لا توجد علاقة بينهما، وبالتالي فحدث ظهور وجه في الرمية الأولى وحدث ظهور وجه في الرمية الثانية هما حدثان مستقلان.

يلزمنا في دراسة الاحتمالات تحديد عدد عناصر فراغ العينة وعدد عناصر المجموعة الجزئية الممثلة للحدث، وفي بعض التجارب يحتوي فراغ العينة وكذلك المجموعة الجزئية الممثلة للحدث على عدد كبير جداً من العناصر، ليس من السهل كتابتها كلها وعدها، لذلك نستخدم بعض طرق العد التي تساعدنا في الحصول على

عدد عناصر فراغ العينة وعدد عناصر المجموعة الجزئية للحدث دون كتابتها، ومن أهم هذه الطرق ما يلي:

(1-5) طرق العد:

(1-5-1) قاعدة الضرب:

إذا كانت التجربة تتم في مرحلتين أو عمليتين وكان عدد النتائج التي نحصل عليها في المرحلة أو العملية الأولى تساوي n_1، وعدد النتائج التي نحصل عليها في المرحلة أو العملية الثانية تساوي n_2، فإن عدد النتائج التي نحصل عليها في المرحلتين أو العمليتين معاً، أي عدد النتائج الكلية لهذه التجربة تساوي n_1n_2.

ويمكن تعميم هذه القاعدة، فإذا كانت التجربة تتم في k من العمليات، وكان عدد النتائج في هذه العمليات $n_1,n_2,n_3,...,n_K$ فإن عدد النتائج الكلية لهذه التجربة يساوي :

$$n_1,n_2.............n_K$$

مثال (1-15):

عند إلقاء قطعتي نقود، نجد أن فراغ العينة لهذه التجربة:

$$S=\{ HH , HT , TH , TT \}$$

فنلاحظ أن عدد النتائج الكلية يساوي 4، وباستخدام قاعدة الضرب نستطيع تحديد عدد النتائج الكلية لهذه التجربة بدون كتابة فراغ العينة، فنجد أن لعملية إلقاء القطعة النقدية الأولى نتيجتان، أي $n_1=2$، ولعملية إلقاء القطعة النقدية الثانية نتيجتان أي $n_2 = 2$، إذن العدد الكلي للنتائج التي نحصل عليها من إلقاء القطعتين يساوي:

$$n_1 n_2 = (2) (2) = 4$$

مثال (1-16):

عند إلقاء مكعبي نرد، ما هو العدد الكلي للنتائج الممكنة؟

الحل:

العدد الكلي للنتائج التي نحصل عليها من المكعب الأول ($n_1 = 6$).

العدد الكلي للنتائج التي نحصل عليها من المكعب الثاني ($n_2 = 6$).

إذن باستخدام قاعدة الضرب نجد أن العدد الكلي للنتائج التي نحصل عليها من تجربة إلقاء مكعبي نرد يساوي:

$$n_1 n_2 = (6)(6) = 36$$

مثال (1-17):

عند إلقاء مكعب نرد وقطعة نقود معـًا، ما هو العدد الكلي للنتائج الممكنة؟

الحل:

العدد الكلي للنتائج التي نحصل عليها من مكعب النرد ($n_1 = 6$).

العدد الكلي للنتائج التي نحصل عليها من قطعة النقود ($n_2 = 2$).

إذن باستخدام قاعدة الضرب نجد أن العدد الكلي للنتائج التي نحصل عليها من إلقاء مكعب نرد وقطعة نقود يساوي:

$$n_1 n_2 = (6)(2) = 12$$

(1-5-2) التباديل:

يمكن تعريف تباديل مجموعة من العناصر المختلفة بأنها عدد الطرق المختلفة التي يمكن أن نرتب بها هذه العناصر ويرمز له بالرمز P_r^n

حيث $$P_r^n = \frac{n!}{(n-r)!}$$

مثال (1-18):

اشتريت مرجعاً من 5 أجزاء. وعلى رف من رفوف مكتبك لا يتوفر إلا 3 أمكنة. بكم طريقة مختلفة يمكنك شغل هذه الأماكن الثلاثة المتوفرة بثلاثة أجزاء تختارها من الأجزاء الخمسة؟

الحل:

عدد الطرائق المختلفة لشغل الأماكن الثلاثة هو عدد متبادلات خمسة أشياء مأخوذ ثلاثة منها في وقت واحد وهو $P_{r}^{n} = P_{3}^{5}$

حيث $P_{3}^{5} = \frac{5!}{(5-3)!} = \frac{5!}{2!} = \frac{(5)(4)(3)(2)(1)}{(2)(1)} = \frac{120}{2} = 60$

أو $P_{3}^{5} = (5)(4)(3) = 60$

مثال (1-19):

كم عدداً مكونـاً من 3 أرقام يمكن تكوينه من الأرقام من 1 إلى 10 مع عدم السماح بالتكرار؟

الحل:

واضح أنه في أي عدد (عينة) يراد تكوينه لا بد من مراعاة الترتيب في أرقامه حيث أن هذا العدد يتغير تبعاً لهذا الترتيب فمثلاً العدد 312 يختلف عن العدد 321 كما أن السحب يتم دون إرجاع نظراً لعدم السماح بتكرار الرقم. أي أن:

$$P_{r}^{n} = P_{3}^{10} = \frac{10!}{(10-3)!} = \frac{10!}{7!} = \frac{(10)(9)(8)(7)(6)(5)(4)(3)(2)(1)}{(7)(6)(5)(4)(3)(2)(1)} = \frac{3628800}{5040} = 720$$

أو $P_{r}^{n} = P_{3}^{10} = \frac{10!}{(10-3)!} = \frac{10!}{7!} = \frac{(10)(9)(8)(7!)}{7!} = (10)(9)(8) = 720$

نتيجة: **يسمى عدد تباديل n من العناصر المميزة المأخوذة سوية هو:**

$$^{n}P_{n} = n(n-1)(n-2)\ldots\ldots\ldots1 = n!$$

والفكرة الأساسية لهذه القاعدة تعتمد على قاعدة الضرب فعند ترتيب هذه العناصر سيكون لدينا **n** من العناصر من الممكن وضع واحدة منها في الترتيب الأول، وسيكون لدينا **(n-1)** من العناصر من الممكن وضع واحدة منها في الترتيب الثاني لأن واحدة من المفردات قد وضعت في الترتيب الأول، وهكذا .... إلى أن نصل إلى الموضع الأخير وهو في هذه الحالة الموضع الخاص بالترتيب **n**، ويكون لدينا خيار واحد فقط لإجراء هذه العملية، لأن **(n-1)** من المفردات قد تم وضعها في المواضع التي سبقتها.

وبتطبيق قاعدة الضرب نجد أن عدد الطرق المختلفة لترتيب **n** من المفردات هو:

$$n\,(n-1)\,(n-2)\ldots\ldots\ldots1$$

وهذا العدد يطلق عليه مضروب n، ويرمز له بالرمز **n!**، مع العلم بأن:

$$1! = 1 \qquad 0! = 1$$

مثال (1-20):

إذا طلبنا من شخص ما إن يرتب 3 كتب (C , B , A) في رف ، فكل الطرق الممكنة في هذه الحالة موضحة فيما يلي :

S= { ABC , ACB ,BAC , BCA , CAB , CBA }

أي توجد 6 طرق لترتيب 3 كتب على رف، وعدد هذه الطرق نستطيع تحديده باستخدام قاعدة التباديل، حيث $(n = 3)$، وذلك كما يلي:

$$P_3^3 = 3! = 3\times2\times1 = 6$$

وبنفس القاعدة نستطيع بسهولة تحديد العدد الكلي للطرق التي يمكن أن نرتب بها أي عدد من المفردات دون كتابة فراغ العينة.

مثال (1-21):

ما العدد الكلي للطرق التي نستطيع بها أن نرتب 7 أطفال في صف؟

الحل:

باستخدام قاعدة التباديل، حيث $(n=7)$، نجد أن العدد الكلي للطرق يساوي:

$$7!=7\times6\times5\times4\times3\times2\times1=5040$$

(1-5-3) التوافيق:

التوافيق هي عدد الطرق التي يتم بها اختيار r من العناصر من بين مجموعة تحتوي على n من العناصر، بحيث أن $(n \geq r)$ مع إهمال عملية الترتيب.

ونرمز للتوافيق بالرمز C_r^n ويحسب كما يلي:

$$\boxed{C_r^n = \frac{n!}{r!\,(n-r)!}}$$

مثال (1-22):

إذا كان لدينا 3 موظفين: أحمد، يوسف، إبراهيم، وأردنا أن نختار منهم لجنة تتكون من موظفين، فما عدد الطرق الممكنة لاختيار هذه اللجنة؟

الحل:

الطرق الممكنة لاختيار هذه اللجنة هي:

$S =$ { أحمد ويوسف، أحمد وإبراهيم ، يوسف وإبراهيم }

أي يوجد **3** طرق لتكوين هذه اللجنة، وبالطبع هنا عملية الترتيب ليس لها أية أهمية، فمثلاً اللجنة المتكونة من أحمد ويوسف، هي نفسها اللجنة المتكونة من يوسف وأحمد.

وكما تلاحظ أننا حددنا عدد الطرق التي يمكن أن نكوِّن بها اللجنة، بكتابة كل الطرق الممكنة ثم عدّها، ولكن هذه الطريقة من الصعب إتباعها عندما يكون العدد الكلي للموظفين كبيراً. فالأسهل هو استخدام قاعدة التوافيق، حيث في هذه الحالة **(n = 3)** ، **(r=2)**، وبالتالي عدد طرق اختيار شخصين من **3** أشخاص مع إهمال الترتيب هو:

$$C_2^3 = \frac{3!}{2!\,(3-2)!} = \frac{3 \times 2 \times 1}{(2 \times 1)(1)} = 3$$

مثال (1-23):

مدرسة بها **6** مدرسات و**3** مدرسين، نريد اختيار لجنة تتكون من **4** أعضـــــــاء.

أ) كم عدد اللجان المختلفة الممكن اختيارها؟

ب) كم عدد اللجان المختلفة الممكن اختيارها بحيث تحتوي اللجنة على **3** مدرسات ومدرس واحد؟

الحل:

أ. عدد اللجان المختلفة الممكن اختيارها، أي عدد اللجان المختلفة التي تحتوي على **4** أشخاص مختارين من **9** أشخاص (**6**+**3**)، هو:

$$C^{9}_{4} = \frac{9!}{4!(9-4)!} = \frac{9\times8\times7\times6\times5!}{4!\times5!} = 126$$

أي سيكون لدينا **126** لجنة كل منها تختلف عن الأخرى (مع إهمال الترتيب).

ب. عدد اللجان المختلفة الممكن اختيارها، بحيث تحتوي اللجنة على **3** مدرسات ومدرس واحد، هو عدد الطرق التي يتم بها اختيار **3** مدرسات من **6** مدرسات واختيار مدرس واحد من **3** مدرسين، أي كأن الاختيار يتم على مرحلتين، فبتطبيق قاعدة الضرب وقاعدة التوافيق، نجد أن عدد الطرق التي تحقق الحدث المطلوب هو:

$$C^{6}_{3} \times C^{3}_{1} = \frac{6!}{3!\times3!} \times \frac{3!}{1!\times2!} = 20\times3 = 60$$

تمارين (1-1)

1. أكتب فراغ العينة للتجارب التالية:

أ) إلقاء مكعب نرد وقطعة نقدية معًا مرة واحدة

ب) إلقاء قطعة نقدية ثلاث مـــرات.

جـ) إذا أخترنا أسرة واحدة عشوائيًا من الأسر التي لديها **4** أطفال، وسألناها عن كل طفل من أطفالها هل هو ذكر أم أنثى.

د) إذا أخترنا ثلاث نساء عشوائيًا، وسألنا كل واحدة منهن هل تستخدم مسحوق تايد لغسل الملابس أم لا.

2. في تجربة إلقاء قطعة نقدية ثلاث مرات، أكتب الفئات الجزئية التي تمثل الأحداث التالية، مع ذكر نوع الحدث:

أ) الحصول على ظهريين أو أكثر.

ب) الحصول على ثلاثة أوجـــه.

ج-الحصول على وجه في الرمية الأولى.

د-الحصول على ثلاثة أوجه أو أقل.

هـ -الحصول على أربعة أوجه

3. في تجربة إلقاء قطعتي نقود معًا، ما هو الحدث المكمل لحدث الحصول على وجهين على الأقل (وجهين أو أكثر) ؟

4. في تجربة إلقاء مكعبي نرد، ما هو الحدث المكمل لحدث الحصول على مجموع يساوي **8** أو اقل؟

5. إذا ألقينا مكعبي نرد، وكان الحدث **A** هو الحصول على مجموع أقل من **5**، والحدث **B** هو الحصول على مجموع أكبر من **10**، والحدث **C** هو الحصول على نفس العدد على المكعبين. فأي حدثين متنافيين وأيهما غير متنافيين؟

6. توجد **4** طرق تربط المدينة **أ**، بالمدينة **ب**، ويوجد طريقان يربطان المدينة ب بالمدينة جـ، فما عدد الطرق التي يمكن أن يسلكها المسافر من **أ**، إلى **جـ مارا**، بالمدينة **ب**؟

7. بكم طريقة يمكننا أن نرتب **5** كتب مختلفة في رف؟

8. بكم طريقة يمكن أن يجلس أستاذ **و3** تلاميذ في صف، بحيث:

أ- يمكن أن يجلس الأستاذ في أي مقعد.

ب- يجب أن يجلس الأستاذ في المقعد الأول.

9. صندوق به **10** كرات (**4 حمراء و 6 بيضاء**)، يراد اختيار **5** كرات معاً من هذا الصندوق، فأحسب ما يلي:

أ- عدد الطرق التي يمكن أن نختار بها من هذا الصندوق **5** كرات.

ب- عدد الطرق التي يمكن أن نختار بها من هذا الصندوق كرة حمراء، و**4** كرات بيضاء.

(1−6) طرق حساب الاحتمالات:

الاحتمال هو مقياس عددي يعبر عن مقدار ثقتنا في إمكانية ظهور حدث ما غير مؤكد الحدوث عند إجراء تجربة معينة. وتوجد طريقتان لحساب الاحتمال، هما الطريقة التقليدية (الطريقة الكلاسيكية)، والطريقة التجريبية (طريقة التكرار النسبي).

(1−6−1) الطريقة التقليدية:

تستخدم هذه الطريقة عندما تكون كل نتائج التجربة لها نفس فرصة الظهور، ويحسب احتمال حدث ما، بقسمة عدد النتائج التي تحقق الحدث على عدد النتائج الكلية للتجربة، أي بقسمة عدد عناصر الفئة الجزئية للحدث على عدد عناصر فراغ العينة.

فإذا رمزنا لعدد عناصر فراغ العينة بالرمز $n\ (S)$، ورمزنا لعدد عناصر الفئة الجزئية التي تمثل الحدث A بالرمز $n\ (A)$، فإن احتمال الحدث A، ويرمز له بالرمز $P(A)$ يحسب كما يلي :

$$\mathbf{P(A)} = \frac{\text{عدد النتائج التي تحقق الحدث } \mathbf{A}}{\text{عدد النتائج الكلية للتجربة}} = \frac{n(A)}{n(S)}$$

ولا نستطيع حساب الاحتمال التقليدي لأي حدث إلا إذا توفر الشرطان التاليان:

1. إذا كانت نتائج التجربة لها نفس فرصة الظهور.
2. أن يكون فراغ العينة للتجربة العشوائية محدود، أي نستطيع أن نحدد عدد عناصره **n(S)**.

مثال (1-24):

عند إلقاء مكعب نرد مرة وحدة واعتبار أن الحدث **A** هو ظهور عدد أكبر من **4** ، فأحسب احتمال **A**.

الحل:

فراغ العينة لهذه التجربــة: $S = \{1, 2, 3, 4, 5, 6\}$

والفئة الجزئية التي تمثل الحدث **A** هي: $A = \{5, 6\}$

$n(A) = 2$ ، $n(S) = 6$

وبالتالي فإن:

احتمال **A** (احتمال ظهور عدد أكبر من **4**) هو:

$$P(A) = \frac{n(A)}{n(S)} = \frac{2}{6}$$

مثال (1-25):

إذا ألقينا قطعة نقود غير متحيزة، فما احتمال ظهور وجه؟

الحل:

فراغ العينة لهذه التجربة: $S = \{H, T\}$

نفرض أن الحدث **A** هو ظهور الوجه، أي أن:

$$A = \{H\}$$

$n(A) = 1$ ، $n(S) = 2$

واحتمال ظهور وجه هو: $P(A) = \frac{n(A)}{n(S)} = \frac{1}{2} = 0.5$

مثال (1-26):

إذا اخترنا عشوائيا أسرة من الأسر التي لديها **3** أطفال، وسألنا الأسرة عن كل طفل من أطفالها هل هو ذكر أم أنثى وافترضنا أن فرصة أن يكون المولود ذكراً تساوي فرصة أن يكون المولود أنثى، فأحسب احتمال أن لدى الأسرة بنتين.

الحل :

1. إذا رمزنا للطفل الذكر بالحرف **b**، وللطفل الأنثى بالحرف ***g***، فإن فراغ العينة لهذه التجربة كما يلي:

$$S = \{bbb, bbg, bgb, gbb, bgg, gbg, ggb, ggg\}$$

والمقصود بالنتيجة ***bbb*** أن الطفل الأول ولد والطفل الثاني ولد والطفل الثالث ولد، والنتيجة ***bbg*** تعني الطفل الأول ولد والطفل الثاني ولد والطفل الثالث بنت، وهكذا ...

إذا فرضنا أن الحدث **A** هو أن للأسرة بنتين، فإن الحدث **A** تحققه النتائج الموضحة فيما يلي:

$$A = [bgg . gbg . ggb]$$

$$n(A) = 3 \qquad ،n(S) = 8$$

وبالتالي فإن الاحتمال المطلوب هو:

$$P(A) = \frac{n(A)}{n(S)} = \frac{3}{8}$$

(1-6-2) الطريقة التجريبية (طريقة التكرار النسبي) :

عند إجراء تجربة بالفعل وتكرار ها تحت نفس الظروف عدداً من المرات، فيعرف التكرار النسبي لحدث معين بأنه عدد المرات التي يظهر فيها هذا الحدث

مقسوماً على العدد الكلي لمرات تكرار التجربة. والاحتمال التجريبي لوقوع حدث معين هو عبارة عن التكرار النسبي لظهوره، وذلك عند تكرار إجراء التجربة تحت نفس الظروف عدداً كبيراً من المرات.

إن الاحتمال المحسوب بطريقة التكرار النسبي قد يتغير عند زيادة تكرار التجربة وحصولنا على معلومات جديدة بخصوص الحدث المطلوب، ولكن هذا التغيير يكون بسيطاً كلما زاد العدد الكلي لمرات تكرار التجربة. ويجب الانتباه أن الاحتمال التجريبي هو تقريب للاحتمال الحقيقي الذي نحصل عليه من الطريقة التقليدية، ويزداد قرباً من الاحتمال الحقيقي كلما زاد عدد مرات إجراء التجربة.

إذا كررنا تجربة ما عدد **n** من المرات تحت نفس الظروف، وكان عدد مرات ظهور الحدث **A** هو **m** من المرات، فيعرف الاحتمال التجريبي للحدث **A** كما يلــي:

الاحتمال التجريبي للحدث **A** = (عدد مرات ظهور الحدث **A**) / (العدد الكلي لمرات إجراء التجربة) $= \frac{m}{n}$

مثال (1-27):

ألقى شخص ما قطعة نقود غير متحيزة **100** مرة، وكان عدد مرات ظهور الوجه **48** مرة، فإن الاحتمال التجريبي لظهور الوجه هو:

$$\frac{48}{100} = 0.48$$

وعندما زاد عدد مرات تكرار التجربة إلى **150** مرة، كان العدد الكلي لظهور الوجه **73** مرة، وأصبح الاحتمال التجريبي لظهور الوجه هو:

$$\frac{73}{150} = 0.4866$$

ونلاحظ أنه كلما زاد عدد المرات الكلية لإجراء التجربة، كلما أقترب الاحتمال التجريبي من **0.50**، وهو احتمال الحصول على وجه عند إلقاء قطعة نقود إذا استخدمنا الطريقة التقليدية كما في مثال **(1–25)**. أي أنه إذا رمينا قطعة النقود عدداً كبيراً من المرات، فإن التكرار النسبي سوف يؤول إلى الاحتمال الحقيقي الذي نحصل عليه من الطريقة التقليدية. ويجب الانتباه أن الاحتمال التجريبي نحصل عليه بعد إجراء التجربة فعلاً، بينما الاحتمال التقليدي يحسب بدون إجراء التجربة.

(1–7) مسلمات الاحتمال:

نستطيع أن نعرف أي احتمال رياضيـاً، بأنه دالة نطاقها فراغ العينة ومداها فئة الأعداد الحقيقية من **0** إلى **1**. وأي احتمال يجب أن يتمتع بالمسلمات التالية:

1. احتمال أي حدث A يجب أن تكون قيمته في المدى من 0 إلى 1 .أي أن :

$$\mathbf{0 \le P(A) \le 1}$$

فأي احتمال لا تزيد قيمته عن **1**، ولا تقل عن **0**، وسنفسر هذه المسلمة في ضوء التعريف التقليدي للاحتمال فيما يلي:

بما أن:
$$\mathrm{P(A)} = \frac{n(A)}{n(S)}$$

حيث: $n(S)$: عـــدد عناصـــر فـــراغ العينـــــة .

$n(A)$: عـدد عناصـر الفئـة الجزئيـة التـي تمثـل الحـدث **A** .

وبما أن عدد عناصر الفئة الجزئية $n(\mathrm{A})$ لأي حدث يجب أن يكون محصوراً بين **0** و $n(\mathrm{S})$، أي أن أقل قيمة يأخذها $n(\mathrm{A})$ هي الصفر، وذلك عندما لا تحتوي الفئة

الجزئية التي تمثل الحدث على أي عنصر، وأكبر قيمة يأخذها هي $n(S)$ وذلك عندما تحتوي الفئة الجزئية التي تمثل الحدث على كل عناصر فراغ العينة، وبالتالي فإن:

$$0 \leq n(A) \leq n\ (s)$$

وبقسمة الأطراف الثلاثة لهذه المتباينة على $n(S)$ نحصل على :

$$\frac{0}{n(S)} \leq \frac{n(A)}{n(S)} \leq \frac{n(S)}{n(S)}$$

$$0 \leq P(A) \leq 1$$

2. احتمال الحدث المؤكد يساوي واحد، واحتمال الحدث المستحيل يساوي صفر، أي أن:

$P(\emptyset) = 0$ $P(S) = 1$

وتفسير هذه المسلمة كما يلي:

إذا كان الحدث **A** حدثاً مؤكداً، فإن:

A = S **وبالتالي فإن** **n(A) = n(S)**

$$P(A) = P(S) = \frac{n(S)}{n(S)} = 1$$

وإذا كان الحدث **A** حدثاً مستحيلاً، فإن:

A= ∅ **وبالتالي فإن:** **n(A) = n(∅)**

$$P(A) = P(\emptyset) = \frac{n(\emptyset)}{n(S)} = \frac{0}{n(S)} = 0$$

3. إذا كان الحدثان A , B حدثين متنافيين ، أي الفئة الجزئية التي تمثل الحدث A منفصلة عن الفئة الجزئية التي تمثل الحدث B فإن :

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

وتفسير هذه المسلمة كما يلي:

بما أن:

$$P(A) = \frac{n(A)}{n(S)} \qquad P(B) = \frac{n(B)}{n(S)}$$

$$P(A \cup B) = \frac{n(A \cup B)}{n(S)} =$$

وبما أن الحدثين B.A منفصلان، أي لا توجد عناصر مشتركة بينهما، فإن عدد العناصر الموجودة في مجموعة الاتحاد تساوي عدد العناصر الموجودة في المجموعة **A** مضافاً إليها عدد العناصر الموجودة في المجموعة **B**، أي أن:

$$P(A \cup B) = \frac{n(A \cup B)}{n(S)} = \frac{n(A) + n(B)}{n(S)} = \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} = P(A) + P(B)$$

وهذه المسلمة يمكن تعميمها إلى أكثر من حدثين متنافيين، فمثلاً إذا كانت الأحداث A_4 , A_3 , A_2, A_1 هي أحداث متنافية فإن:

$$P(A_1 U A_2 U A_3 U A_4) = P(A_1) + P(A_2) + P(A_3) + P(A_4)$$

وهكذا ....

تمارين (1-2)

1. إذا ألقينا قطعتي نقود معـًا، فأحسب احتمال ظهور:

أ- نتائج متشابهة على القطعتين.

ب- وجه أو أكثر.

جـ- **3** أوجه.

2. إذا ألقينا مكعب نرد، فأحسب احتمال ظهور:

أ- عدد أكبر من **3**.

ب- عدد محصور بين **1** و **6**.

جـ- عدد أقل من **8**.

د- ظهور العدد **8**.

3. إذا ألقينا **3** قطع نقدية معـًا، فأحسب احتمال الحصول على:

أ- نتائج متشابهة على القطع الثلاث.

ب- وجهين أو أكثر.

جـ- وجهين أو أقـل.

4. إذا ألقينا مكعبي نرد معـًا، فأحسب ما يلي:

أ- احتمال الحصول على مجموع يسـاوي **11**.

ب- احتمال الحصول على مجموع يساوي **9** أو أكثر.

جـ- احتمال أن يظهر العدد **3** على أحد المكعبين.

5. إذا ألقينا قطعتي نقود ومكعب نرد معـًا، فما احتمال الحصول على وجهين على قطعتي النقود وعدد أكبر من **4** على مكعب النرد؟

6. إذا اخترنا عشوائياً عائلة واحدة من العائلات التي لديها **4** أطفال، فأحسب الاحتمالات التالية بافتراض أن فرصة أن يكون المولود ذكر تساوي فرصة أن يكون المولود أنثى:

أ- احتمال أن يكون الطفل الأول ولد.

ب- احتمال أن يكون عدد الإناث في العائلة يساوي عدد الذكور.

7. تم اختيار مجموعة تشمل **200** وحدة منتجة من سلعة ما، ووجدنا **24** وحدة منها تالفة، فإذا اخترنا عشوائياً مفردة من هذه المجموعة فما احتمال أن تكون تالفة؟

8. تم اختيار مجموعة تشمل **1200** رجلاً من سكان مدينة ما، ووجدنا عدد المدخنين في هذه المجموعة يساوي **720** رجلاً، فإذا اخترنا عشوائياً رجلاً من هذه المجموعة فما احتمال أن يكون مدخناً؟

9. أشترك محمد وعلي في لعبة، وكرروها **25** مرة، فاز محمد في **10** مرات، وفاز علي في **8** مرات، وتعادلاً في الباقي، فإذا لعبا محمد وعلي هذه اللعبة فأحسب ما يلي:

أ- احتمال فوز محمـــــد.

ب- احتمال عدم فوز علي.

11. ما هي القيم التي لا تمثل احتمالات؟ ولماذا؟

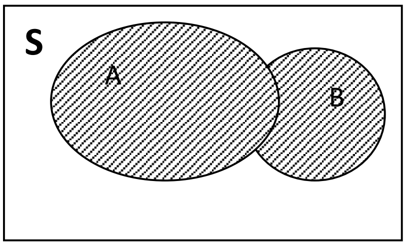
القيم هي: -0.40، 3.0، 1.75 ، 0.20، -1.0، 1.0، $\frac{20}{19}$ ، $\frac{5}{3}$ ، $\frac{2}{8}$

(1–8) قانون جمع الاحتمالات:

قانون جمع الاحتمالات هو القانون الذي نحسب منه احتمال اتحاد حدثين أو أكثر، وسنتعرض فيما يلي لقانون جمع الاحتمالات في حالة حدثين غير متنافيين، وفي حالة حدثين متنافيين.

(1–8–1) قانون الجمع لحدثين غير متنافيين:

إذا كان الحدثان **A , B** غير متنافيين أي أن الحدثين **A , B** يمكن أن يحدثا معاً في نفس الوقت، فإن احتمال حدوث الحدث $(A \cup B)$ يعني احتمال حدوث **A** أو **B** أو الاثنين معاً، أي ظهور أحد الحدثين على الاقل، وهذا الحدث يمثله الأجزاء المظللة في شكل (1–4).



شكل (1–4)

واحتمال حدوث الحدث $(A \cup B)$ يساوي عدد العناصر التي تحتويها مجموعة الاتحاد $A \cup B$ مقسوماً على عدد عناصر فراغ العينة، وبما أن عدد عناصر مجموعة الاتحاد يساوي عدد العناصر التي تحتويها **A** مضافاً إليه عدد العناصر التي تحتويها **B** مطروحاً منه عدد العناصر التي تحتويها مجموعة التقاطع $A \cap B$، لأن عناصر مجموعة التقاطع $n(A \cap B)$ قد قمنا بجمعها مرتين، جمعناها مرة مع

المجموعة A ومرة أخرى مع المجموعة B ، ولذا لابد من طرحها مرة واحدة حتى نحصل على المجموع الصحيح ، أي أن :

$$P(A \cup B) = \frac{n(A \cup B)}{n(S)} = \frac{n(A) + n(B) - n(A \cap B)}{n(S)}$$

$$= \frac{n(A)}{n(S)} + \frac{n(B)}{n(S)} - \frac{n(A \cap B)}{n(S)} = P(A) + P(B) - P(A \cap B)$$

وبالتالي احتمال اتحاد أي حدثين غير متنافيين A , B يحسب من القانون التالي والذي يطلق عليه قانون جمع الاحتمالات لحدثين غير متنافيين :

$$\boxed{P(A \cup B) = P(A) + P(B) - P(A \cap B)}$$

حيث:

$P(A \cup B)$: احتمال ظهور الحدث A أو الحدث B أو الاثنين معـًا. أي احتمال ظهور أحد الحدثين على الأقل.

P (A) : احتمال ظهور الحدث A .

P (B) : احتمال ظهور الحدث B .

$P(A \cap B)$: احتمال ظهور الحدثين A و B معـًا في نفس الوقت.

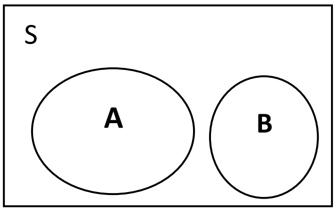
(1-8-2) قانون الجمع لحدثين متنافيين:

إذا كان الحدثان B, A متنافيين أي أن وقوع أحدهما يمنع وقوع الأخر، أي لا توجد نتائج مشتركة بينهما، أي لا توجد نتائج تحقق الحدث A وتحقق الحدث B في

نفس الوقت ، كما هو واضح في شكل (1−5)، وبذلك استحالة حدوثهما معـًا، أي أن:

$$\mathbf{A \cap B = \emptyset}$$

$$\mathbf{P(A \cap B) = 0}$$



شكل (1−5)

وبالتالي عندما يكون الحدثان **A،B** متنافيين، فإن احتمال حدوث الحدث $\boldsymbol{A \cup B}$، يعني احتمال ظهور الحدث A أو ظهور الحدث B، ويحسب كما يلي :

$$\boxed{\mathbf{P(A \cup B) = P(A) + P(B)}}$$

لاحظ أننا في هذه الحالة لا نستعمل عبارة (أو الاثنين معـًا) لأن الحدثين متنافيان ولا يمكن الحصول عليهما معـًا.

كما يجب الانتباه أن وجود **أو** في الاحتمال المطلوب، يعني الاحتمال المطلوب هو احتمال فئة الاتحاد، ونستطيع حسابه من قانون جمع الاحتمالات.

مثال (1-26):

عند إلقاء مكعب نرد مرة واحدة ما هو احتمال الحصول على رقم فردي أو رقم أكبر من **3**؟

الحل:

فراغ العينة لهذه التجربة: **$S = \{1, 2, 3, 4, 5, 6\}$**

لنفرض أن:

الحدث **A** يمثل حدث ظهور رقم فــــردي.

الحدث **B** يمثل حدث ظهور رقم أكبر من **3**.

$$A = \{1, 3, 5\} \quad P(A) = \frac{3}{6}$$

$$B = \{4, 5, 6\} \qquad P(B) = \frac{3}{6}$$

وبما أن: $A \cap B = \{5\} \quad P(A \cap B) = \frac{1}{6}$

وبما أن مجموعة التقاطع ليست مجموعة خالية، إذن الحدثان غير متنافيين؛

ولحساب الاحتمال المطلوب نستخدم القانون التالي:

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

$$= \frac{3}{6} + \frac{3}{6} - \frac{1}{6} = \frac{5}{6}$$

مثال (1-27):

إذا علمت أن احتمال نجاح طالب ما في مادة الإحصاء يساوي **0.70**، واحتمال أن ينجح في مادة الرياضة **0.65**، واحتمال أن ينجح في المادتين معاً **0.52**، فأحسب احتمال أن ينجح في إحدى المادتين على الأقل.

الحل:

احتمال أن ينجح الطالب في إحدى المادتين على الأقل، يعني احتمال أن ينجح في الإحصاء أو أن ينجح في الرياضة أو أن ينجح في المادتين معاً، فهنا حدث النجاح في الإحصاء وحدث النجاح في الرياضة غير متنافيين. فإذا فرضنا أن:

الحدث **A** يمثل النجاح في الإحصاء.

الحدث **B** يمثل النجاح في الرياضـــة.

فالاحتمال المطلوب هو: $P(A\cup B) = ?$

والاحتمـــــالات المعطـــــاة هـــــي:

$P(A)=0.70$, $P(B)= 0.65$, $P(A\cap B) = 0.52$

وحيث أن الحدثين **A** ، **B** غير متنافيين، فإن الاحتمال المطلوب هو :

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

$$= 0.70 + 0.65 - 0.52 = 0.83$$

مثال)1-28(:

عند إلقاء مكعب نرد مرة واحدة ما هو احتمال ظهور عدد زوجي أو **5** ؟

الحل:

فراغ العينة لهذه التجربة: $S=\{1,2,3,4,5,6\}$

لنفرض أن:

الحدث **A** يمثل حدث ظهور عدد زوجي.

الحدث **B** يمثل حدث ظهور العـــــدد **5**.

$$A = \{2,4,6\} \quad P(A) = \frac{3}{6}$$

$$B = \{5\} \quad P(B) = \frac{1}{6}$$

$$(A\cap B) = \{\ \} \quad P(A\cap B) = 0$$

وبما أن مجموعة التقاطع مجموعة خالية، إذن الحدثان متنافيان، ولحساب الاحتمال المطلوب نستخدم القانون التالي:

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

$$= \frac{3}{6} + \frac{1}{6} = \frac{4}{6} = \frac{2}{3}$$

(1–9) الاحتمال الشرطي:

إذا كان الحدثان **B,A** حدثين غير مستقلين، أي أن ظهور أحدهما يؤثر في احتمال ظهور الأخر. فالاحتمال الشرطي هو احتمال ظهور أحدهما وليكن **B** ، إذا علمت أن الحدث الآخر وليكن **A** قد ظهر فعلاً. ويرمز لهذا الاحتمال بالرمز **P(B/A)**، ويقرأ احتمال ظهور الحدث **B** بشرط أن الحدث **A** قد ظهر فعلاً. أو احتمال ظهور الحدث **B** إذا علمت أن الحدث **A** قد ظهر فعلاً. ويحسب كما يلي:

$$P(B/A) = \frac{P(A\cap B)}{P(A)}$$

حيث: $P(A) > 0$

أما الاحتمال الشرطي **P(A/B)** ، يقرأ احتمال ظهور الحدث **A** بشرط أن الحدث **B** قد ظهر فعلاً. أو احتمال ظهور الحدث **A** إذا علمت أن الحدث **B** قد ظهر فعلاً، ويحسب كما يلي:

$$P\,(A/B) = \frac{P(A\cap B)}{P(B)}$$

حيث: $P(B) > 0$

لاحظ أن الحدث الذي ظهر فعلاً أي تم وقوعه فعلاً، يكتب على يمين الشرطة، أي على يمين العلامة (/)، واحتماله يكون موجود في المقام. وذلك كما هو واضح في العلاقتين السابقتين.

أما إذا كان الحدثان **B , A** حدثين مستقلين ، أي أنه:

إذا كان ظهور الحدث **A** لا يؤثر على احتمال ظهور الحدث **B** فإن:

$$P(B/A) = P(B)$$

وإذا كان ظهور الحدث **B** لا يؤثر على احتمال ظهور الحدث **A** فإن:

$$P(A/B) = P(A)$$

مثال (1-29):

إذا ألقينا مكعب نرد مرة واحدة، ما هو احتمال ظهور عدد زوجي إذا علمت أن العدد الظاهر أكبر من **4**؟

الحل:

فراغ العينة لهذه التجربة: **S= {1,2,3,4,5,6}**

لنفرض أن:

الحدث **A** يمثل حدث ظهــور عدد زوجـــي.

الحدث **B** يمثل حدث ظهور عدد أكبر مــن **4**.

كما ذكرنا سابقـًا أن الحدث الذي ظهر فعلاً يوضع على يمين العلامة (/)،

وبما أن الحدث الذي علمنا أنه ظهر أعطيناه الرمز **B**، إذن الاحتمال المطلوب هو

P (A/B) ويحسب كما يلي :

$$P(A/B) = \frac{P(A\cap B)}{P(B)}$$

بما أن: $A = \{2, 4, 6\}$ $\quad B = \{5, 6\}$, $A\cap B = \{6\}$

إذن: $P(A\cap B) = \frac{1}{6}$ $\quad P(B) = \frac{2}{6}$

$$P(A/B) = \frac{P(A\cap B)}{P(B)} = \frac{1/6}{2/6} = \frac{1}{2}$$

مثال (1-30):

مدرسة بها **50** مدرسة مصنفه حسب الجنسية والمستوى التعليمي كما يلي:

الجنسية / المستوى التعليمي	ليبية	غير ليبية
خريجة معهد	18	6
خريجة جامعة	22	4

اخترنا منهن مدرَّسة واحدة عشوائيًا.

أ- إذا علمت أنها ليبية، فما احتمال أن تكون خريجة جامعة؟

ب- إذا علمت أنها خريجة جامعة، فما احتمال أن تكون ليبية؟

الحل:

نفرض أن:

الحدث **A** هو أن تكون المدرَّسة ليبية.

الحدث **B** هو أن تكون المدرَّسة خريجة جامعة .

أ- الاحتمال المطلوب في (أ)، هو P (B/A) ، ويحسب كما يلي :

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

حيث P (A) هو احتمال أن تكون المدرَّسة ليبية، وهو يساوي عدد المدرسات الليبيات مقسوماً على العدد الكلي للمدرسات، أي يساوي :

$$P(A) = \frac{40}{50}$$

و$P(A \cap B)$ هو احتمال أن تكون المدرسة ليبية وخريجة جامعة في نفس الوقت، وهو يساوي عدد المدرسات الليبيات الخريجات من الجامعة مقسوماً على العدد الكلي للمدرسات ، أي يساوي :

$$P(A \cap B) = \frac{22}{50}$$

إذن احتمال أن تكون خريجة جامعة مع العلم بأنها ليبية هو:

$$P(B/A) = \frac{P(A \cap B)}{P(A)} = \frac{22/50}{40/50} = \frac{22}{40} = 0.55$$

ب- الاحتمال المطلوب في (ب)، هو $P(A/B)$، ويحسب كما يلي :

$$P(A/B) = \frac{P(A \cap B)}{P(B)}$$

حيث P(B) هو احتمال أن تكون المدرَّسة خريجة جامعة، وهو يساوي عدد المدرسات خريجات الجامعة مقسوماً على العدد الكلي للمدرسات، أي:

$$P(B) = \frac{26}{50}$$

إذن احتمال أن تكون ليبية مع العلم بأنها خريجة جامعة هو:

$$P(A/B) = \frac{P(A \cap B)}{\mathrm{P(B)}} = \frac{22/50}{26/50} = \frac{22}{26} = 0.8461$$

مثال (1-31):

صندوق به **6** كرات بيضاء **و4** كرات حمراء، فإذا سحبنا منه كرتين عشوائيًا، وعملت أن الكرة الأولى بيضاء، فما احتمال أن تكون الثانية حمراء؟

1. مع عدم ترجيع الكرة الأولى قبل سحب الكرة الثانية.
2. مع ترجيع الكرة الأولى قبل سحب الكرة الثانيــــــة.

الحل:

نفــرض أن:

الحدث **A** هو أن تكون الكرة الأولى بيضاء.

الحدث **B** هو أن تكون الكرة الثانية حمراء.

الصندوق في هذه التجربة يعتبر كأنه فراغ العينة، لأنه يحتوي على كل الكرات التي يمكن أن تظهر لنا منها كرة (أي يحتوي على كل النتائج الممكنة).

الاحتمال المطلوب هو **P(B/A)** ، وهو احتمال أن تكون الكرة الثانية حمراء مع العلم بأن الكرة الأولى التي سحبت من الصندوق كانت بيضاء.

أ. في حالة **عدم ترجيع** الكرة الأولى قبل سحب الكرة الثانية، لتحديد هذا الاحتمال يجب معرفة ما تبقى في الصندوق بعد عملية السحب الأولى، فهنا الحدثان **B,A** حدثان **غير مستقلين**، وذلك لأننا لم نرجع الكرة المسحوبة في المرة الأولى إلى الصندوق، مما يؤدي إلى تغير محتوياته ، وبالتالي يتأثر احتمال ظهور كرة حمراء. وحيث أن الباقي في الصندوق **9** كرات هي: **5** بيضاء **و4** حمراء. فاحتمال أن تكون

الكرة المسحوبة في المرة الثانية حمراء مع العلم أن الأولى بيضاء يساوي عدد الكرات الحمراء المتبقية في الصندوق مقسوماً على العدد الكلي للكرات المتبقية، أي الاحتمال المطلوب هو:

$$P(B/A) = \frac{4}{9}$$

ب. في حالة **ترجيع** الكرة الأولى قبل سحب الثانية، يكون الحدثان **مستقلين**، أي أن ظهور كرة بيضاء في السحبة الأولى لا يؤثر في احتمال أن تكون الثانية حمراء، لأننا رجَّعنا الكرة الأولى قبل سحب الثانية، وبالتالي فمحتويات الصندوق لم تتغير، فكأننا لم نسحب منه أي كرة ولذلك فإن:

احتمال أن تكون الثانية حمراء $= \frac{4}{10}$

ملاحظة:

يجب الانتباه إلى أن اختيار الأشخاص دائماً مع عدم الإرجاع، حتى إذا لم يذُكر ذلك صراحة. فليس من المنطق أن نعيد نفس الأسئلة على نفس الشخص أكثر من مرة.

(1–10) قانون ضرب الاحتمالات:

إذا كان الحدثان **B, A** أي حدثين في فراغ العينة **S**، فإننا نستطيع الحصول على $P(A \cap B)$، أي احتمال ظهور **A** و **B** معاً في نفس الوقت، كما يلي:

(1–10–1) قانون الضرب لحدثين غير مستقلين:

علمنا أنه عندما يكون الحدثان **B , A** غير مستقلين، والحدث **A** هو الذي تم ظهوره أولاً، فإن الاحتمال الشرطي يحسب كما يلي :

$$P(B/A)=\frac{P(A\cap B)}{\mathrm{P(A)}} \quad ، \qquad \mathrm{P(A)}>0$$

ومـن هــذه الصيغــة نستنتــج الصيغــة الخاصــة بحســـاب

فنجد أن: $P(A\cap B)= \mathrm{P}(A).\ P\ (B/A) \qquad P\ (A)>0$

وبالتالي نجد انه إذا كان الحدث **A** هو الذي تم ظهوره أولاً، فإن احتمال الحصول على الحدثين **A** و**B** معاً في نفس الوقت يحسب من الصيغة التالية التي يطلق عليها قانون ضرب الاحتمالات لحدثين غير مستقلين:

$$\boxed{P(A\cap B)=\mathrm{P}(A).P\ (B/A) \quad P\ (A)>0}$$

(1-10-2) قانون الضرب لحدثين مستقلين:

أما إذا كان الحدثان **B , A** مستقلين ، فقد علمنا انه:

$$P(B/A)\ =P(B)$$

وبالتالي تصبح صيغة قانون ضرب الاحتمالات عندما يكون الحدثان **B ,A** مستقلين كما يلي:

$$\boxed{\mathrm{P}(\mathrm{A}\cap B)\ =P(A).P(B)}$$

ملاحظة:

الحرف **و** مقرونـاً بمفهوم التقاطع، فالحدث الذي يقصد به ظهور **A** و **B** معـًا، يعني ظهور الحدث $\mathrm{A}\cap B$، ونستطيع حساب احتمال حدث من هذا النوع، بتطبيق التعريف التقليدي للاحتمال مباشرة، أي بقسمة عدد عناصر مجموعة التقاطع $\mathrm{n(A}\cap B)$ على عدد عناصر فراغ العينة **n(S)**، أو باستخدام قانون ضرب الاحتمالات.

مثال (1-32):

صندوق به **5** كرات بيضاء و**10** كرات خضراء، فإذا سحبنا منه كرتين عشوائيًا، فما احتمال أن تكون الأولى خضراء والثانية بيضاء؟

أ. في حالة السحب مع عدم الإرجـــاع.

ب. في حالـــة السحـــب مـــع الإرجـــاع.

الحل:

نفرض أن:

الحدث **A** هو أن تكون الكرة الأولى خضراء.

الحدث **B** هو أن تكون الكرة الثانية بيضـــــاء.

الاحتمال المطلوب هو:

$$\mathbf{P(A \cap B) = ?}$$

أ- **في** حالة أنّ السحب تم مع عدم الإرجاع، تكون الأحداث غير مستقلة، ويحسب الاحتمال المطلوب كما يلي:

$$\mathbf{P(A\cap B) = P(A).P(B/A)}$$
$$\mathbf{= \frac{10}{15} \times \frac{5}{14} = \frac{5}{21}}$$

ب- **في** حالة أنَّ السحب تم مع الإرجاع، تكون الأحداث مستقلة، ويحسب الاحتمال المطلوب كما يلي:

$$\mathbf{P(A\cap B) = P(A)\,.P(B)}$$
$$\mathbf{= \frac{10}{15} \times \frac{5}{15} = \frac{2}{9}}$$

مثال)1-33(:

صندوق به **6** كرات بيضاء و**14** كرة سوداء، فإذا سحبنا منه كرتين عشوائيًا، فما احتمال أن تكون الكرتان من نفس اللون؟

أ. في حالة السحب مع عدم الإرجاع.

ب. في حالــة السحــب مع الإرجاع.

الحل:

الاحتمال المطلوب هو احتمال أن تكون الكرتان من نفس اللون، أي احتمال أن تكون الكرتان بيضاء أو أن تكون الكرتان سوداء.

نفرض أن:

الحدث W_1 هو أن تكون الكرة الأولى بيضــاء.

الحدث W_2 هو أن تكون الكرة الثانية بيضـــاء.

الحدث B_1 هو أن تكون الكرة الأولى ســوداء.

الحدث B_2 هو أن تكون الكرة الثانية ســـوداء.

الاحتمال المطلوب هو:

احتمال (W_1 و W_2) **أو** (B_1 و B_2)، وكما ذكرنا سابقًا أن حرف و يعني تقاطع وحرف **أو** يعني اتحاد، إذن الاحتمال المطلوب هو:

$$P\left((W_1 \cap W_2) \cup (B_1 \cap B_2)\right) = ?$$

وبما أن حدث الحصول على كرتين بيضاء وحدث الحصول على كرتين سوداء، هما حدثان متنافيان، لأنهما لا يمكن أن يظهرا معـًا، وباستخدام قانون جمع الاحتمالات لحدثين متنافيين يكون الاحتمال المطلوب مساويـًا ما يلي:

$$P\left((W_1 \cap W_2) \cup (B_1 \cap B_2)\right) = P(w_1 \cap w_2) + P(B_1 \cap B_2)$$

وبتطبيق قانون الضرب نحصل على الاحتمال المطلوب، حيث:

أ- الاحتمال المطلوب في حالة السحب مع عدم الإرجاع:

$$P((w_1 \cap w_2) \cup (B_1 \cap B_2)) = P(W_1 \cap W_2) + P(B_1 \cap B_2)$$

$$= P(w_1)P(W_2/W_1) + P(B_1)P(B_2/B_1)$$

$$= \left(\frac{6}{20} \times \frac{5}{19}\right) + \left(\frac{14}{20} \times \frac{13}{19}\right) = \frac{30+182}{380} = 0.5578$$

ب- الاحتمال المطلوب في حالة السحب مع الإرجاع:

$$P((w_1 \cap w_2) \cup (B_1 \cap B_2)) = P(W_1 \cap W_2) + P(B_1 \cap B_2)$$

$$= P(W_1)P(W_2) + P(B_1)P(B_2)$$

$$= \left(\frac{6}{20} \times \frac{6}{20}\right) + \left(\frac{14}{20} \times \frac{14}{20}\right) = \frac{36+19}{400} = 0.5800$$

فيما يلي نعرض بعض الأمثلة التي نستعين لحلها بطرق العد المذكورة في بداية هذا الفصل.

مثال (1-34):

قفل خزنة يتكون من **4** خانات، ولكي يفتح يجب أن يُحدد في كل خانة عدد من الأعداد الصحيحة من **0** إلى **9**، فما احتمال سرقة هذه الخزنة؟

الحل:

سنستعين لحل هذا المثال بقاعدة الضرب، فهنا لدينا **4** خانات أي **4** مراحل وكل خانة ممكن أن نضع فيها أي عدد من الأعداد الصحيحة **0**، **1**،...، **9**، أي كل مرحلة ممكن إنجازها بعشرة طرق، وبالتالي فإن $n_1 = 10$ ، $n_2 = 10$ ، $n_3 = 10$ ، $n_4 = 10$ ويكون العدد الكلي للحالات التي يمكن أن نضبط عليها خانات قفل الخزنة (عدد عناصر فراغ العينة) يساوي:

$$n_1n_2n_3n_4 = (10)\,(10)\,(10)\,(10) = 10000$$

أي توجد **10000** حالة يمكن أن نضبط عليها خانات قفل هذه الخزنة، وبالطبع حالة واحدة فقط من هذه الحالات هي التي تفتح قفل الخزنة (الحالة التي يستعملها صاحب الخزنة)، إذن احتمال سرقة هذه الخزنة أي احتمال فتحها هو:

$$\frac{1}{10000} = 0.0001$$

مثال (1-35):

إذا طلبنا من أحد الأشخاص أن يرتب جلوس مدير و 3 مدرسين و 3 طلبة في مقاعد مرقمة من **1** إلى فأحسب ما يلي:

أ- احتمال أن يجلس المدير والمدرسين الثلاثة في المقاعد الأربعة الأولى، ثم الطلبة في المقاعد الباقية.

ب- احتمال ان يجلس المدير في المقعد الأول ثم يجلس المدرسين الثلاثة في المقاعد الثاني والثالث والرابع، ثم يجلس الطلبة في بقية المقاعد.

الحل:

واضح في هذا المثال أننا **نهتم بالترتيب**، وبالتالي طريقة العد التي سنستعين بها لحل هذا المثال، هي **قاعدة التباديل**. فيكون:

عدد الطرق الكلية التي يمكن أن نرتب بها **7** أشخاص في صف = **!7**

أ. الحدث المطلوب احتماله هو أن يجلس المدير والمدرسين الثلاثة في المقاعد الأربعة الأولى، ويجلس الطلبة في بقية المقاعد. فهنا كأن الحدث يتم على مرحلتين:

المرحلة الأولى: يقوم فيها بترتيب المدير والمدرسين الثلاثة في المقاعد الأربعة الأولى ويتم ذلك بعدد من الطرق = **!4**.

المرحلة الثانية: يقوم فيها بترتيب الطلبة الثلاثة في المقاعد الثلاثة الباقية، ويتم ذلك بعدد من الطرق = **!3**.

وبتطبيق قاعدة الضرب نحصل على عدد الطرق التي تحقق الحدث المطلوب (عدد عناصر الفئة الجزئية للحدث) وهو = **(!3) (!4)**.

إذن الاحتمال المطلوب يساوي:

$$\frac{(4!)\ (3!)}{7!} = \frac{1}{35}$$

ب. الحدث المطلوب احتماله هو ان يجلس المدير في المقعد الأول والمدرسين الثلاثة في المقاعد الثاني والثالث والرابع، ويجلس الطلبة في بقية المقاعد، فهنا كأن الحدث يتم على **3** مراحل:

المرحلة الأولى: يقوم فيها بوضع المدير في المقعد الأول، ويتم ذلك بعدد من الطرق = **!1**.

المرحلة الثانية: يقوم فيها بترتيب المدرسين الثلاثة في المقعد الثاني والثالث والرابع، ويتم ذلك بعدد من الطر ق = **!3**.

المرحلة الثالثة: يقوم فيها بترتيب الطلبة الثلاثة في المقاعد الثلاثة الباقية، ويتم ذلك بعدد من الطرق = !3

وبتطبيق قاعدة الضرب نحصل على عدد الطرق التي تحقق الحدث المطلوب (عدد عناصر الفئة الجزئية للحدث) وهو = (!3)(!3)(!1).

إذن الاحتمال المطلوب يساوي:

$$\frac{(1!)(3!)(3!)}{7!} = \frac{1}{140}$$

مثال (1-36):

فصل دراسي يحتوي على 8 طالبات و7 طلبة، فإذا اخترنا عشوائياً من هذا الفصل مجموعة تتكون من 5 أشخاص (**طبعاً الاختيار تم مع عدم الإرجاع**)، فما احتمال أن تحتوي هذه المجموعة على 3 طالبات وطالبين؟

الحل:

الاحتمال المطلوب هو احتمال اختيار مجموعة تتكون من 3 طالبات وطالبين، وبالطبع نهمل الترتيب، وهذا الاحتمال يساوي عدد النتائج التي تحقق هذا الحدث مقسوماً على العدد الكلي للنتائج الممكنة.

واضح في هذا المثال أننا نهتم **بالاختيار مع اهمال الترتيب**، وبالتالي طريقة العد التي سنستعين بها لحل هذا المثال هي **قاعدة التوافيق** فيكون:

العدد الكلي للنتائج الممكنة (عدد عناصر فراغ العينة) يساوي العدد الكلي للطرق التي يمكن أن نختار بها 4 اشخاص من 15 شخص (8+7)، أي يساوي:

$$C_5^{15} = \frac{15!}{5!\,(15-5)!} = \frac{15 \times 14 \times 13 \times 12 \times 11 \times 10!}{5! \times 10!} = 3003$$

وعدد الطرق التي تحقق الحدث المطلوب يساوي عدد الطرق التي يتم بها اختيار **3** طالبات من **8** طالبات واختيار طالبين من **7** طلبة، أي كأن الاختيار يتم على مرحلتين، وبالتالي فبتطبيق قاعدة الضرب وقاعدة التوافيق، نجد أن عدد الطرق التي تحقق الحدث المطلوب هو:

$$C_3^8 \quad C_2^7 = \frac{8!}{3! \times 5!} \times \frac{7!}{2! \times 5!} = 1176$$

إذن الاحتمال المطلوب هو: $= \frac{1176}{3003} = 0.3916$

ونستطيع حساب الاحتمال المطلوب في خطوة واحدة كما يلي:

الاحتمال المطلوب هو: $= \frac{C_3^8 \times C_2^7}{C_5^{15}} = 0.3916$

ملاحظة:

عندما تتم عملية السحب العشوائي للمفردات بدون إرجاع، نستطيع حساب أي احتمال مطلوب باستخدام قوانين الاحتمالات أو باستخدام قاعدة التوافيق ونحصل على نفس النتيجة تماماً. لأنه عند إيجاد التوافيق لا نختار نفس المفردة مرتين، أي كأن السحب تم بدون إرجاع.

فمثلاً نستطيع حساب الاحتمال المطلوب في (أ) الخاص بالمثال **(1–33)** باستخدام قاعدة التوافيق، وذلك كما يلي:

العدد الكلي للنتائج الممكنة (عدد عناصر فراغ العينة) يساوي العدد الكلي للطرق التي يمكن أن نختار بها كرتين من **20** كرة بالصندوق **(6+14)**، أي يساوي:

$$C_2^{20} = \frac{20!}{2!(20-2)!} = \frac{20!}{2! \times 18!} = 190$$

ويتحقق الحدث المطلوب عندما تكون الكرتان بيضاء أو عندما تكون الكرتان سوداء، وهما حدثان متنافيان، وبالتالي:

احتمال أن تكون الكرتان من نفس اللون يساوي:

احتمال أن تكون الكرتان بيضاوين + احتمال أن تكون الكرتان سوداوين.

فبتطبيق قاعدة التوافيق، نستطيع أن نحسب كلا من هذين الاحتمالين كما يلـــــــي:

احتمال أن تكون الكرتان بيضاء يساوي عدد الطرق التي يمكن أن نختار بها كرتين بيضاء من **6** كرات بيضاء بالصندوق مقسوماً على العد الكلي لاختيار كرتين من **20** كرة بالصندوق، أي يساوي:

$$\frac{C_2^6}{C_2^{20}} = \frac{15}{190}$$

واحتمال أن تكون الكرتان سوداوين يساوي عدد الطرق التي يمكن أن نختار بها كرتين سوداوين من **14** كرات سوداء بالصندوق مقسوماً على العد الكلي لاختيار كرتين من **20** كرة بالصندوق، أي يساوي:

$$\frac{C_2^{14}}{C_2^{20}} = \frac{91}{190}$$

إذن الاحتمال المطلوب $0.5578 = \frac{92}{190} + \frac{91}{190} =$

وهي نفس النتيجة التي حصلنا عليها بحل المثال عن طريق قانون الضرب في حالة عدم الإرجاع. وبالتالي في حالة الاختيار مع عدم الإرجاع يستطيع الطالب أن يستخدم قانون الضرب أو التوافيق، مع العلم بأن قاعدة التوافيق أسهل بكثير عندما يزيد عدد المفردات المختارة على أثنين.

ملخص الفصل الأول

تعرضنا في هذا الفصل لتعريف مصطلحات هامة وهي التجربة العشوائية، فراغ العينة، الحدث وأنواعه، ثم تطرقنا إلى الطرق التي تساعدنا في معرفة عدد عناصر فراغ العينة وعدد عناصر الفئة الجزئية التي تمثل الحدث وهي (قاعدة الضرب، التباديل، التوافيق)، ثم درسنا كيفية حساب الاحتمال بالطريقة التقليدية والطريقة التجريبية، حيث الاحتمال التقليدي للحدث **A** يحسب كما يلي:

$$\mathrm{P(A)} = \frac{\text{عدد النتائج التي تحقق الحدث } A}{\text{عدد النتائج الكلية للتجربة}} = \frac{n(\mathrm{A})}{n(S)}$$

أما الاحتمال التجريبي للحدث **A** يحسب كما يلي :

$$\boldsymbol{\frac{m}{n}} = \frac{\text{عدد مرات ظهور الحدث } \boldsymbol{A}}{\text{العدد الكلي لمرات إجراء التجربة}} = \text{الاحتمال التجريبي للحدث } \boldsymbol{A}$$

تم عرضنا القوانين المستخدمة لحساب الاحتمالات وهي:

قانون جمع الاحتمالات ويستخدم للحصول على احتمال اتحاد حدثين، حيث:

1- عندما يكون الحدثان غير متنافيين $\mathbf{P(A\cup B) = P(A) + P(B) - P(A \cap B)}$

2- عندما يكون الحدثان متنافيين $\mathbf{P(A\cup B) = P(A) + P(B}$

قانون الاحتمال الشرطي $\mathbf{P(B/A)}$، وهو احتمال ظهور الحدث **B** بشرط أن الحدث **A** قد ظهر فعلاً . ويحسب كما يلي:

$$P(B/A) = \frac{P(A \cap B)}{P(A)}$$

قانون ضرب الاحتمالات، ويستخدم لحساب احتمال تقاطع حدثين، حيث:

1. عندما يكون الحدثان غير مستقلين $P(A \cap B) = P(A).P(B/A)$
2. عندما يكون الحدثان متنافيـــــــن $P(A \cap B) = P(A).P(B)$

تمارين (1-3)

1. إذا ألقينا مكعب نرد، فاحسب احتمال الحصول على:

أ- عدد زوجي أو عدد أكبر من **4**.

ب- عدد فردي أو العـــــــدد **6**.

2. إذا ألقينا 3 قطع نقدية معـًا، فاحسب ما يلي:

أ- احتمال الحصول على نتائج متشابهة أو **3** وجوه.

ب-احتمال الحصول على وجهين أو ظهريـــــــن.

3. إذا ألقينا مكعبي نرد معـًا، فأحسب ما يلي:

أ- احتمال الحصول على مجموع أكبر من **10** أو نتائج متشابهة على المكعبين.

ب- احتمال الحصول على مجموع يساوي **8** أو مجموع يساوي **12**.

4. إذا علمت أن: **P(A) = 0.45**، **P(B) = 0.65**، واحتمال الحصول على الحدثين معـًا = **0.20**، فاحسب احتمال ظهور أحد الحدثين على الأقل.

5. إذا علمت أن احتمال أن ينجح طالب ما في مادة الإحصاء يساوي **0.70**، واحتمال أن ينجح في مادة الرياضية **0.62**، واحتمال أن ينجح في إحدى المادتين على الأقل **0.89** فأحسب احتمال أن ينجح في المادتيـــن معـًا.

6. إذا ألقي مكعب نرد، وعلمت أنه تم الحصول على عدد زوجي فما احتمال أن يكون هذا العدد أكبر من **2**؟

7. تُستخدم حافلتان لنقل موظفي شركة معينة، فإذا علمت أن احتمال أن تكون الحافلة الأولى عاطلة عن العمل **0.05** واحتمال أن تكون الحافلة الثانية عاطلة عن العمل **0.02**، فأحسب احتمال أن تكون الحافلتان عاطلتين عن العمل.

8. صندوق **به 4** كرات صفراء و **6** كرات بيضاء. إذا سحبنا من هذا الصندوق كرتين، فما احتمال أن تكون الأولى صفراء والثانية بيضاء؟

أ- إذا تم السحب مع الإرجــــــاع.

ب -إذا تم السحب مع عدم الإرجاع.

9. صندوق **به 7** كرات حمراءو **5** كرات بيضاء. إذا سحبنا من هذا الصندوق كرتين، ما احتمال ان تكونا من نفس اللـــــون.

أ- إذا تم السحب مع الإرجـــــاع.

ب-إذا تم السحب مع عدم الإرجاع.

11. حقيبة سفر قفلها يتكون من **3** خانات، ولكي تفتح يجب أن يُحدد في كل خانة عدد صحيح من **0** إلى **9**، فما احتمال سرقة أغراض من هذه الحقيبة.

11. تتكون أسرة من أب وأم **3** بنات و**4** أولاد، جلسوا عشوائياً في صف به **9** مقاعد، فأحسب الاحتمالات التالية:

أ-أن يجلس الاب في المقعد الأول، ويجلس بقية أفراد الاسرة في المقاعد الأخرى.

ب-أن يجلس الأب والأم في المقعدين الأولين ويجلس بقية أفراد الأسرة في المقاعد الأخرى.

جـ- أن يجلس الأب في المقعد الأول وتجلس الأم في المقعد الثاني، ويجلس بقية أفراد الأسرة في بقية المقاعد.

د- أن يجلس الأب في المقعد الأول وتجلس الأم في المقعد الأخير ويجلس أطفالهم بينهما.

12. صندوق يحتوي **10** وحدات منتجة **(8** جيدة الصنع و**2** بها تلفْ**)**، فإذا سحبنا من هذا الصندوق **4** وحدات معًا، فأحسب ما يلي:

أ−احتمال ظهور **3** وحدات جيدة ووحدة تالفـــــة؟

ب−احتمال أن تكون كل الوحدات المسحوبة جيـدة؟

الفصل الثاني
المتغيرات العشوائية وتوزيعاتها الاحتمالية

(2-1) المتغير العشوائي:

المقصود بالمتغير بصفة عامة هو الخاصية أو الظاهرة محل الدراسة والبحث، وقد أطلق عليها مصطلح متغير لأن قيمتها تتغير من مفردة إلى أخرى، فمثلاُ إذا كانت دراستنا خاصة بأطوال طلبة المرحلة الثانوية، ففي هذه الدراسة تكون المفردة هي الطالب، والخاصية المستهدفة بالدراسة هي الطول، وبما أن الطول يتغير من مفردة إلى أخرى أي من طالب إلى أخر فيسمى متغير.

وإذا كانت التجربة التي نجريها هي تجربة عشوائية، فسيكون اختيارنا للمفردة عشوائيـًا، أي خاضع لعامل الصدفة دون تدخل العامل البشري فيها، وبما أن المفردة هي التي تحدد قيمة المتغير (**في المثال السابق، الطالب هو الذي يحدد قيمة الطول**) فإن المتغير يُسمى في هذه الحالة المتغير العشوائي.

كذلك في بعض التجارب العشوائية، لا تهمنا النتيجة التي نحصل عليها من التجربة في حد ذاتها، ولكن ينصب اهتمامنا على متغير معين ترتبط قيمته بكل نتيجة من نتائج التجربة العشوائية، فمثلاً عندما تكون التجربة العشوائية هي إلقاء 3 قطع نقدية ويكون اهتمامنا منصبّـًا على عدد الأوجه التي نحصل عليها، فهنا عدد الأوجه هو متغير عشوائي لأنه يتغير من نتيجة إلى أخرى وقيمته التي سنحصل عليها تحددها النتيجة التي نحصل عليها من التجربة، وبما أن ظهور أي نتيجة عشوائي، إذن نطلق على عدد الأوجه التي نحصل عليها عند إلقاء 3 قطع نقدية مصطلح المتغير العشوائي وبصفة عامة نستطيع تعريف المتغير العشوائي كما يلي:

المتغير العشوائي:

هو متغير كمي نعبر عنه في تحديد قيمه على كل نتيجة من النتائج التي يمكن أن نحصل عليها من إجراء تجربة عشوائية أي على كل عنصر من عناصر فراغ عينة تجربة عشوائية، وبالتالي فإن المتغير العشوائي هو دالة نطاقها فراغ العينة و مداها هو فئة الأعداد الحقيقية.

وأية قيمة من قيم المتغير العشوائي قد تحددها نتيجة واحدة أو أكثر من نتائج فراغ العينة. ولكن كل نتيجة من نتائج فراغ العينة تقابلها قيمة واحدة فقط للمتغير العشوائي.

وكما وضحّنا نستخدم كلمة عشوائي للدلالة على أن القيم التي يمكن أن يأخذها المتغير تعتمد على نتائج تجربة عشوائية. ويرمز عادة للمتغير العشوائي بأحد الحروف **X , Y , Z** .....

ويمكن تصنيف المتغيرات العشوائية إلى نوعين:

1. متغيرات عشوائية متقطعة (منفصلة).
2. متغيرات عشوائية مستمرة (متصلة).

وفيما يلي سنتعرض لكل نوع من هذين النوعين:

(2-1-1) المتغيرات العشوائية المتقطعة (المنفصلة):

يعرف المتغير العشوائي المتقطع كما يلي:

تعريف المتغير العشوائي المتقطع:
هو متغير عشوائي قيمه التي يمكن أن يأخذها منفصلة عن بعضها وقابلة للعدّ.

مثال (2-1):

عند إلقاء مكعب نرد مرة واحدة، فإن فراغ العينة كما عرفنا هو:

$$S=\{1,2,3,4,5,6\}$$

فإذا كان المتغير العشوائي **X** يمثل عدد النقاط التي نحصل عليها عند إلقاء مكعب نرد، سنجد أن القيم التي يمكن أن يأخذها هذا المتغير العشوائي هي:

1 ، 2 ، 3 ، 4 ،5 ، 6

فنلاحظ أن هذه القيم منفصلة عن بعضها فالمتغير يمكن أن يأخذ القيمة **1** أو القيمة **2** ولكن لا يمكن أن يأخذ القيم التي بين هاتين القيمتين، فمثلاً لا يمكنه أن يأخذ القيمة **1.67**، وكذلك يمكنه أن يأخذ القيمة **2** والقيمة **3** ولكن لا يمكنه أن يأخذ القيم التي بينهما، وهكذا ..... فنجد أن القيم التي يمكن أن يأخذها هذا المتغير منفصلة عن بعضها ولذلك يسمى بالمتغير المنفصل أو المتقطع. وبما أن القيم منفصلة عن بعضها فتكون قابلة للعد، فنجد أن عدد القيم التي يمكن أن يأخذها هذا المتغير يساوي **6** قيم.

مثال (2-2):

إذا كان المتغير العشوائي **X** يمثل عدد الأوجه التي يمكن أن نحصل عليها عند إلقاء قطعتي نقود ، فإن فراغ العينة لهذه التجربة:

S = (HH ,HT ,TH ,TT)

من فراغ العينة نستطيع تحديد القيم التي يمكن أن يأخذها المتغير **X** وهي:

X = 0 ذلك عند ظهور النتيجة **T T** (عدم الحصول على وجه).

X = 1 وذلك عند ظهور النتيجة **H T** أو **T H** (الحصول على وجه واحد).

X = 2 وذلك عند ظهور النتيجة **H H** (الحصول على وجهين).

فإن المتغير العشوائي **X** يمكنه أن يأخذ **3** قيم وهي القيم **0** ، **1** ، **2** ، وبما أن قيم المتغير منفصلة وقابلة للعد، إذن فالمتغير الذي يمثل عدد الأوجه التي نحصل عليها عند إلقاء قطعتي نقود هو متغير متقطع.

وبصفة عامة فكل المتغيرات التي نحصل على قيمها عن طريق العد هي متغيرات متقطعة (منفصلة).

(2-1-2) المتغيرات العشوائية المستمرة (المتصلة):

يعرف المتغير العشوائي المستمر كما يلي:

تعريف المتغير العشوائي المستمر (المتصل):
هو المتغير العشوائي الذي يمكن أن يأخذ أي قيم في فترة معينة، أي تكون القيم التي يمكن أن يأخذها المتغير العشوائي متصلة ببعضها أي مستمرة في فترة معينة، وبالتالي فهي غير قابلة للعدّ.

مثال (2-3):

إذا كان المتغير العشوائي يمثل أطوال طلبة إحدى الجامعات خلال سنة معينة، فلو فرضنا أن أقصر طالب قامته **155** سم وأطولهم قامة **170** سم، فإذا كانت التجربة العشوائية هي اختيار طالباً واحداً عشوائياً من هذه الجامعة، فهنا المتغير وهو طول الطالب المختار عشوائياً يمكنه أن يأخذ أية قيمة في المدى من **155** سم إلى **170** سم، فمثلاً يمكنه أن يأخذ القيمة **160.02** سم ويمكنه أن يأخذ القيمة **165.731** سم وهكذا ... وبالتالي فإنه يمكن أن يأخذ أية قيمة في الفترة من **155** إلى **170**. وبالتالي فالقيم التي يمكن أن يأخذها هذا المتغير هي قيم متصلة أي مستمرة في الفترة من **155**

إلى **170**، ولذلك يطلق على هذا النوع من المتغيرات اسم المتغيرات المتصلة أو المستمرة.

مثال (2-4) :

مجموعة من الطلبة أوزانهم محصورة ما بين القيمتين **55** كيلو جرام و **105** كيلو جرام، فإذا اخترنا من هؤلاء الطلبة طالباً واحداً عشوائياً، ورمزنا لوزن الطالب المختار بالمتغير العشوائي **X** ، فسنجد أن المتغير العشوائي **X** يمكنه أن يأخذ القيمة **55** أو القيمة **105** أو أية قيمة محصورة بينهما، وبالتالي سنجد أن القيم التي يمكن أن يأخذها هذا المتغير عددها لانهائي ولا يمكن كتابتها كقيم منفصلة نستطيع عدها، فقيمة مستمرة ومتصلة ببعض. فهذا المتغير يمكنه أن يأخذ أية واقعة في الفترة من **55** إلى **105**. وبصفة عامة فكل المتغيرات التي نحصل على قيمها بالقياس هي متغيرات مستمرة.

(2-2) التوزيعات الاحتمالية:

كما علمنا أن المتغيرات العشوائية تنقسم إلى نوعين وهما، المتغيرات العشوائية المتقطعة والمتغيرات العشوائية المستمرة، وبالتالي ستنقسم التوزيعات الاحتمالية هي الأخرى إلى نوعين، فإذا كان التوزيع الاحتمالي خاص بمتغير عشوائي متقطع فيسمي توزيع احتمالي متقطع، وإذا كان التوزيع الاحتمالي خاص بمتغير عشوائي مستمر فيسمي توزيع احتمالي مستمر. وفيما يلي سنتعرض لتعريف كل نوع من هذين النوعين.

(2–2–1) التوزيع الاحتمالي المتقطع (المنفصل).

التوزيع الاحتمالي المتقطع عبارة عن جدول يحتوي على كل القيم التي يمكن ان يأخذها المتغير العشوائي المتقطع مقرونة باحتمالاتها، وأحيانا يعبر عن التوزيع الاحتمالي بصيغة رياضية معينة تسمي **دالة كتلة الاحتمال** ويرمز لها بالرمز.

$f(x)$ وهي تعطي الاحتمالات التي تأخذها القيم المختلفة للمتغير العشوائي المتقطع X .

فإذا كان لدينا المتغير العشوائي المتقطع X فنستطيع التعبير عن التوزيع الاحتمالي بدالة كتلة الاحتمال كالتالي.

$$\boxed{f(x) = P(X = x)}$$

وبالتعويض في دالة كتلة الاحتمال عن اية قيمة من قيم المتغير العشوائي نحصل عن احتمال الحصول على تلك القيمة، أي أن قيمة دالة كتلة الاحتمال هي احتمال، فمثلا $f(5)$ هو احتمال أن يأخذ المتغير العشوائي المتقطع X القيمة 5 أي:

$$f(5) = P(X=5)$$

تعريف التوزيع الاحتمالي المتقطع: هو عبارة عن جدول يشمل كل القيم التي يمكن أن يأخذها المتغير العشوائي المتقطع مقرونة باحتمالاتها، ويعبر عن احتمالات القيم التي يمكن أن يأخذها المتغير العشوائي المتقطع بصيغة رياضية تسمي دالة كتلة الاحتمال.

مثال (2-5):

إذا ألقينا مكعب نرد مرة واحدة، وكان المتغير العشوائي X يمثل العدد الذى يظهر على الوجه. فهنا القيم التي يمكن أن يأخذها هذا المتغير العشوائي هي القيم:

6,5,4,3,2,1

والتوزيع الاحتمالي (دالة كتلة الاحتمال) لهذا المتغير يمثله جدول **(2−1)**.

جدول (2−1)

x	1	2	3	4	5	6
$f(x)$	1/6	1/6	1/6	1/6	1/6	1/6

ويمكن التعبير عن هذا التوزيع الاحتمالي بالصيغة الرياضية $f\ (x)$ والتي يطلق عليها دالة كتلة الاحتمال، حيث :

$$f(x) = \begin{cases} \frac{1}{6} & x = 1, 2, 3, 4, 5, 6 \\ 0 & \textbf{otherwise} \end{cases}$$

حيث المقصود بكلمة (otherwise) لأية قيمة أخرى، أي قيمة دالة كتلة الاحتمال تساوي صفر لأية قيمة أخرى غير القيم المحددة وهي **1،2،3،4،5،6**. ويجب ان يتحقق في أي توزيع احتمالي متقطع (أي في أية دالة كتلة احتمال) الشرطان التاليان:

(1) $0 \leq f(x) \leq 1$ لأن $f(x)$ تمثل احتمال، ونعلم ان أي احتمال يجب ألا يكون سالبا ولا يزيد عن الواحد الصحيح.

(2) $\sum f(x) = 1$ لأن $\sum f(x)$ هو عبارة عن مجموع احتمالات كل القيم التي يمكن ان يأخذها المتغير العشوائي المتقطع والتي تعتمد على كل نتائج فراغ العينة أي أن:

$$\sum f(x) = P(S) = 1$$

شرطي التوزيع الاحتمالي المتقطع:

(1) $0 \leq f(x) \leq 1$ لأي قيمة x.

(2) $\sum f(x) = 1$

من دالة كتلة الاحتمال للمتغير العشوائي المتقطع **X**، نستطيع تحديد احتمال أية قيمة يمكن ان يأخذها هذا المتغير العشوائي **X**، وذلك بالتعويض مباشرة في دالة كتلة الاحتمال بالقيمة المراد حساب احتمال أن يأخذها المتغير العشوائي المتقطــــع **X** .

مثال (2-6):

إذا ألقينا قطعة نقدية واحدة مرتين، وكان المتغير العشوائي **X** يمثل عدد المرات التي نحصل فيها على وجه.

أ. حدد القيم التي يمكن ان يأخذها المتغير العشوائـــــــــــــي **X** .

ب. اوجد التوزيع الاحتمالي لهذا المتغير العشوائـــــــــــــــــــــي.

جـ . عبر عن هذا التوزيع الاحتمالي بصيغة رياضية لدالة الاحتمال $f(x)$.

د. أحسب الاحتمالات التالية:

$$\mathbf{P(x \geq 1), p(x > 1), p(x = 0)}$$

الحـل:

أ. لتحديد القيم التي يأخذها المتغير العشوائي، يجب أولا كتابة فراغ العينة لهذه التجربة العشوائية وهي رمى قطعة نقدية مرتين، حيث:

S = { HH , HT , TH ,TT }

القيم x التي يمكن ان يأخذها المتغير العشوائي موضحة في جدول **(2-2)**

جدول (2-2)

النتيجة	x
HH	2
HT , TH	1
TT	0

إذن القيم التي يمكن أن يأخذها المتغير العشوائي الذي يمثل عدد المرات التي نحصل فيها على وجه عند إلقاء قطعة نقدية مرتين هي: **0،1،2**.

ب. التوزيع الاحتمالي لهذا المتغير العشوائي يوضحه جدول **(2−3)**، حيث احتمال أي قيمة $f(x)$ هو عدد نتائج فراغ العينة المناظرة لهذه القيمة مقسوما على عدد النتائج الكلية (عدد عناصر فراغ العينة).

جدول (2−3)

x	0	1	2
$f\ (x)$	1/4	2/4	1/4

جـ. الصيغة الرياضية لهذا التوزيع الاحتمالي المتقطع، أي دالة كتلة الاحتمال $f(x)$ هي:

$$f(x) = \begin{cases} 1/4 & x = 0\ ,2 \\ 2/4 & x = 1 \\ 0 & \text{otherwise} \end{cases}$$

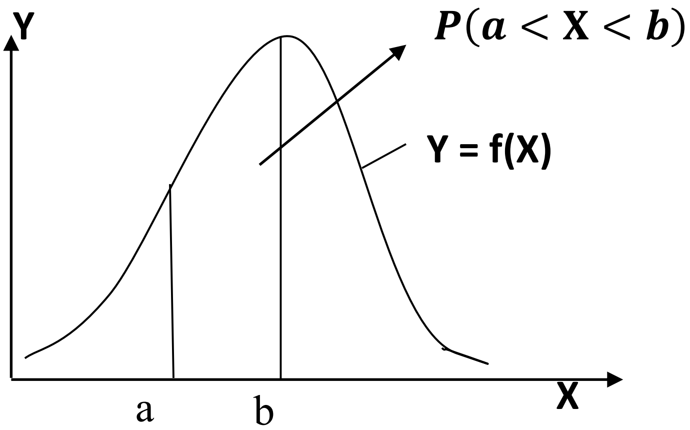
د. الاحتمـــالات المطلوبـــــة:

$$P(X = 0) = f\ (0) = \frac{1}{4}$$
$$P(X > 1) = f\ (2) = \frac{1}{4}$$
$$P(X \geq 1) = f(1) + f\ (2) = \frac{2}{4} + \frac{1}{4} = \frac{3}{4}$$

(2−2−2) التوزيع الاحتمالي المستمر (المتصل):

حيث أن المتغير العشوائي المستمر يتعامل مع فترات وليس مع قيم منفصلة، فلا نستطيع التعبير عن التوزيع الاحتمالي للمتغير العشوائي المستمر بجدول، بينما نعبر عنه بدالة تسمي **دالة كثافة الاحتمال** ويرمز لها كذلك بالرمز $f(x)$. حيث المساحة المحصورة بين منحني هذه الدالة ومحور السينات فوق فترة معينة يساوي احتمال أن يقع المتغير العشوائي المستمر داخل هذه الفترة، فمثلا احتمال أن يأخذ

المتغير المستمر أية قيمة داخل فترة معينة ولتكن الفترة $[a\,,b]$ مساويا للمساحة المحصورة بين محور السينات ومنحنى دالة كثافة الاحتمال لهذه الفترة وذلك كما هو واضح في **شكل (2−1)**.



شكل (2−1)

وبصفة عامة دالة كثافة الاحتمال هي دالة متصلة (مستمرة) ومعرفة عند جميع قيم **x** في المجال (∞،∞−) ويجب أن يتوفر فيها ما يلي:

شرطي دالة كثافة الاحتمال:

(1) $f(x) \geq 0$ لجميع قيم المتغير العشوائي المستمر x

(2) المساحة الكلية المحصورة بين المنحني الذي يمثل دالة كثافة الاحتمال ومحور السينات مساوية الواحد الصحيح.

إذا كانت دالة كثافة الاحتمال دالة خطية، فنستطيع الحصول على المساحة التي تمثل الاحتمال المطلوب من الرسم، لان في هذه الحالة سيكون الشكل المراد الحصول على مساحته شكلا منتظما (مثلثا أو مربعا أو مستطيلا أو شبة منحرف) وكل هذه الأشكال نعرف القوانين التي نحسب منها مساحتها. اما إذا كانت دالة كثافة الاحتمال دالة غير خطية، فنحسب المساحة عن طريق تكامل دالة كثافة

الاحتمال، ولن نتعرض لدوال غير خطية لأن موضوع التكامل خارج موضوع هذا الكتاب.

وحيث أن المتغير المستمر يتعامل مع فترات ولا يتعامل مع قيم معينة، وبالتالي فاحتمال أن يأخذ المتغير العشوائي المستمر $(\mathbf{X})$ قيمة معينة ولتكن $(\boldsymbol{a})$ يساوي دائما صفر، وذلك لأن المساحة فوق قيمة معينة عبارة عن خط، والخط ليس لديه مساحة. وبما ان احتمال ان يساوي المتغير المستمر قيمة معينة يساوي صفر، إذن عندما يكون المتغير $(\mathbf{X})$ متغيرا مستمرا، تكون الاحتمالات الأربعة التالية متساوية.

$$\boxed{\mathbf{P}(\mathbf{a} < \boldsymbol{X} < \boldsymbol{b}) = \mathbf{P}(\mathbf{a} \le \mathbf{X} < \boldsymbol{b}) = \mathbf{P}(\mathbf{a} < \boldsymbol{X} \le \boldsymbol{b}) = \mathbf{P}(\mathbf{a} \le \mathbf{X} \le \mathbf{b})}$$

وهذا يعني أن الفترة التي نحسب احتمالها تشمل الحد الأدني أو الحد الأعلى أوكليهما او لا تشملهما، لا يؤثر ذلك على قيمة الاحتمال في حالة المتغير العشوائي المستمر.

مثال (2-7):

إذا كانت دالة كثافة الاحتمال للمتغير العشوائي المستمر **X** كما يلي:

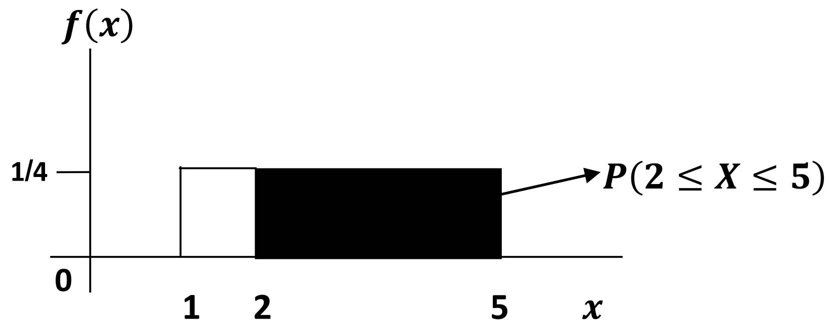
$$\boldsymbol{f}(\boldsymbol{x}) = \begin{cases} \mathbf{1/4} & \mathbf{1} \le \boldsymbol{x} \le \mathbf{5} \\ \mathbf{0} & \boldsymbol{otherwise} \end{cases}$$

1. ارسم دالة كثافة الاحتمال لهذا المتغير.
2. احسب $P\ (2 \ge X \ge 5)$.

الحل:

لرسم دالة الاحتمال، نجعل المحور السيني يمثل القيم التي يمكن أن يأخذها المتغير العشوائي (X)بينما المحور الصادي يمثل قيم دالة كثافة الاحتمال$f(x)$.

وبما ان الدالة خطية، فالصورة البيانية لها عبارة عن خط مستقيم، ونلاحظ أن قيمة الدالة في الفترة من **1** إلى **5** هي قيمة ثابتة تساوي $\frac{1}{4}$ ، وبالتالي فإن الخط الذي يمثل الدالة في هذه الفترة هو خط يوازي محور السينات، مرسومـًا عند $\frac{1}{4}$ =$f(x)$، اما بالنسبة لأي فترة أخرى فسيكون الخط الذي يمثل الدالة منطبقا على محور السينات، لان دالة كثافة الاحتمال تساوي صفر بالنسبة لأي فترة أخرى غير الفترة من **1** إلى **5**. وشكل **(2-2)** يوضح الصورة البيانية لهذه الدالة.



شكل (2-2)

2- الاحتمال المطلوب هو المساحة المحصورة بين المحور الأفقي والخط المستقيم الذي يمثل الدالة وبين المستقيمين x=**2** ، x=**5** . وهي المساحة المظللة في شكل **(2-2)**. ولأن الدالة خطية فنستطيع حساب المساحة من الرسم، فنجد أن المساحة المظللة مستطيل عرضه $\frac{1}{4}$ وطوله **3** إذن:

المساحة المظللة (الاحتمال المطلوب) = العرض ×الطول =$\frac{1}{4}\times 3 = \frac{3}{4}$

مثال)2-8(:

إذا كانت دالة كثافة الاحتمال للمتغير المتصل **X** كما يلي :

$$f(x) = \begin{cases} x/2 & 0 \le x \le 2 \\ 0 & otherwise \end{cases}$$

أ. ارسم دالة كثافة الاحتمال لهذا المتغير.

ب. أحسب $P(0 \le X \le 1)$

الحل

بما ان الدالة خطية، فالصورة البيانية لها هي عبارة عن خط مستقيم، ولرسم هذا الخط المستقيم يلزمنا تحديد الاحداثي $(x\,, f(x))$لاي نقطتين من النقط التي يمر بها، فنجد انه:

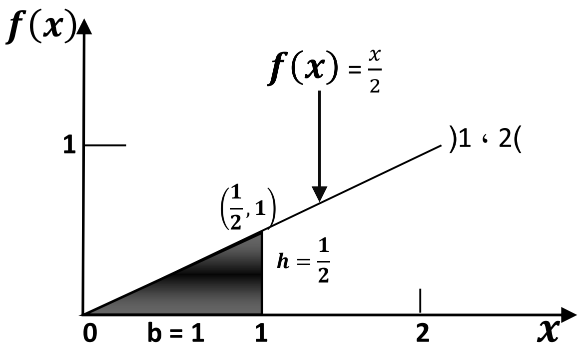
عندما **x=0**، بالتعويض في دالة كثافة الاحتمال نجد أن **0**= **f(0)** **أذن:**

النقطة **(0،0)** تقع على الخط المستقيـــــــــــــــــم.

عندما **x=2**، بالتعويض في دالة كثافة الاحتمال نجد ان **1**= **f(2)** **إذن:**

النقطة **(2،1)** هي كذلك تقع على الخط المستقيم.

بتحديد هاتين النقطتين والتوصيل بينهما نحصل على الخط المستقيم الذي يمثل دالة كثافة الاحتمال في الفترة من **0** إلى **2**، اما بالنسبة لأي فترة اخري فسيكون الخط الذي يمثل الدالة منطبقا على محور السينات لان دالة كثافة الاحتمال تساوي صفر بالنسبة لأي فترة اخرى غير الفترة المذكورة. وشكل **(2،3)** يوضح الصورة البيانية لهذه الدالة.



شكل (2−3)

2.الاحتمال المطلوب هو المساحة المحصورة بين المحور الأفقي والخط المستقيم الذي يمثل الدالة وبين المستقيمين $x = 0$ ، $x = 1$ وهي المساحة المظللة في شكل **(2–3)**. وفي هذه الحالة لان الدالة خطية فنستطيع حساب المساحة من الرسم، فنجد ان المساحة المظللة هي مثلث إبعاده كما يلي:

قاعدته **(b) = 1 – 0= 1**

ارتفاعه **(h) = f (1)= $\frac{1}{2}$**

بما ان مساحة المثلث =**1/2** (القاعدة) × (الارتفاع).

إذن المساحة المظللة (الاحتمال المطلوب)= $(\frac{1}{2})(1)(\frac{1}{2}) = \frac{1}{4}$

تمارين (1-2)

1. اذكر مع التعليل أيا من الجداول التالية يمثل توزيعـًا احتماليا متقطعا، وأيا منها ليس كذلك، مع توضيح السبب:

أ.

x	2	3	4	5
$f(x)$	1/8	3/8	5/8	−1/8

ب.

x	−1	0	1	2
$f(x)$	1/5	4/10	2/10	1/5

جـ.

x	11	13	15	17	19
$f(x)$	1/7	1/7	3/7	2/7	2/7

2. عند إلقاء **3** قطع نقدية، واعتبار أن المتغير العشوائي **X** يمثل عدد الأوجه التي نحصل عليها.

أ. أوجد التوزيع الاحتمالي لهذا المتغير العشوائي.

ب. عبر عن هذا التوزيع الاحتمالي بصيغة رياضية لدالة الاحتمال $f(x)$.

جـ. احسب الاحتمالات التالية:

$$P(0 < X \leq 2) \quad ، \quad P\,(X \geq 2) \quad ، \quad P(X = 1)$$

3. عند إلقاء مكعبي نرد، واعتبار أن المتغير العشوائي **X** يمثل مجموع العددين الظاهرين على المكعبين، أوجد التوزيع الاحتمالي لهذا المتغير.

4. اوجد قيمة الثابت b التي تجعل الجدول التالي يمثل توزيعا احتماليا متقطعا.

x	2	3	4	5	6
$f(x)$	1/8	2/8	3/8	b	1/8

5. إذا كانت دالة الاحتمال للمتغير العشوائي المتقطع **X** كما يلي:

$$f(x) = \begin{cases} 1/k & x = 1, 2, 3, 4, 5 \\ 0 & otherwise \end{cases}$$

أ. أوجد قيمة **k**.

ب. أحسب الاحتمالات التالية:

$P(X = 6)$ ، $P(0 < X \leq 2.5)$ ، $P(X \geq 4)$ ، $P(1 < X \leq 3)$

6. إذا كانت دالة كثافة الاحتمال للمتغير العشوائي المستمر **X** كما يلي:

$$f(x) = \begin{cases} 1/5 & 3 \leq x \leq 8 \\ 0 & otherwise \end{cases}$$

أحسب $P(2 \leq X \leq 7)$.

7. إذا كانت دالة كثافة الاحتمال للمتغير العشوائي المستمر **X** كما يلي :

$$f(x) = \begin{cases} \frac{4 - x}{8} & 0 \leq x \leq 4 \\ 0 & otherwise \end{cases}$$

أحسب $P(X \leq 2)$

(2–3) وصف التوزيعات الاحتمالية:

توجد مقاييس إحصائية لوصف التوزيعات الاحتمالية الخاصة لأى متغير عشوائي سواء كان متقطعا أو مستمرا. وذلك لمعرفة الخواص العامة لتغيرات الظاهرة محل الدراسة التي يمثلها المتغير العشوائي.

وحيث أن التوزيع الاحتمالي يتحدد من كل القيم التي يمكن أن يأخذها المتغير العشوائي والتي تعتمد على كل النتائج الممكن الحصول عليها من تجربة عشوائية، وبالتالي فإن أي توزيع احتمالي لمتغير عشوائي يمثل توزيع احتمالي للمجتمع الذي يتكون من كل القيم التي يمكن أن يأخذها هذا المتغير.

فمثلا إذا كانت دراستنا خاصة بأطوال طلبة كلية الاقتصاد في سنة معينة، فسيتكون المجتمع من أطوال كل طلبة الاقتصاد في السنة المعنية بالدراسة. فإذا رمزنا للظاهرة المستهدفة بالدراسة في هذه الحالة وهي ظاهرة الطول بالرمز **X** فالتوزيع الاحتمالي للمتغير العشوائي **X** يسمي التوزيع الاحتمالي لمجتمع الاطوال.

توزيع المجتمع:
هو التوزيع الاحتمالي للبيانات المجمعة عن كل مفردات المجتمع.

واي مقاييس إحصائية خاصة بالمجتمع ككل كمقاييس النزعة المركزية (المتوسط الحسابي، الوسيط المنوال،...) أو مقاييس التشتت (التباين، الانحراف المعياري ،...) او اية مقاييس إحصائية اخرى تسمي معالم، لأنها تصف لنا المجتمع وتحدد معالمه، فتحسب المعلمة باستخدام بيانات عن كل مفردات المجتمع دون استثناء.

تعريف المعلمة:
هي أي مقياس إحصائي يحسب من كل بيانات المجتمع.

والمعالم عبارة عن قيم ثابتة لا تتغير، لان المجتمع محل الدراسة ثابت لا يتغير أثناء إجراء الدراسة، ولذلك يطلق على المعالم احيانا الثوابت الإحصائية.

وعادة تستخدم الحروف اليونانية للتعبير عن المعالم فيرمز للوسط الحسابي للمجتمع بالحرف μ (ميو) ولتباين المجتمع بالرمز σ^2 (سيجما تربيع) وللانحراف المعياري للمجتمع بالحرف σ (سيجما) وهكذا ....

ومن اهم المقاييس التي يهتم بها علم الإحصاء، هي مقاييس النزعة المركزية ومقاييس التشتت، حيث تعرض الطالب في منهج السنة الثانية لكيفية حسابها من التوزيعات التكرارية، اما في هذا المنهج فإننا سنتعرض لكيفية حساب أهم مقياس من مقاييس النزعة المركزية وهو الوسط الحسابي، واهم مقياسين للتشتت وهما التباين والانحراف المعياري، من التوزيعات الاحتمالية.

(2-3-1) الوسط الحسابي.

إذا كان لدينا مجتمع يتكون من قيم يمثلها المتغير العشوائي المتقطع **X** ، حيث التوزيع الاحتمالي لهذا المجتمع كما يلي:

x	x_1	x_2	....	x_r
$f(x)$	$f(x_1)$	$f(x_2)$	....	$f(x_r)$

فنرمز للوسط الحسابي للمجتمع بالرمز μ، ويحسب باستخدام الصيغة التالية:

$$\mu = \sum_{i=1}^{i=r} x_i \ f(x_i)$$

أي أن الوسط الحسابي لتوزيع احتمالي متقطع هو مجموع حاصل ضرب كل قيمة من القيم التي يمكن أن يأخذها المتغير العشوائي في الاحتمالات المناظرة لتلك القيم.

أما أذا كان التوزيع الاحتمالي مستمرا، فعند حساب الوسط الحسابي نستخدم التكامل بدلا من المجموع، ولكننا لن نتعرض لهذا الموضوع.

مثال (2-9) :

أذا كان لدينا مجتمع يتكون من درجات 10 طلبة وكانت الدرجات كما يلي:

7 ، 6 ، 5 ،4 ، 6 ، 6 ، 7 ، 4 ، 5 ، 5

أوجد التوزيع الاحتمالي لهذا المجتمع، ثم إحسب منه الوسط الحسابي للمجتمع (الوسط الحسابي للدرجات).

الحل:

الظاهرة المستهدفة بالدراسة (المتغير العشوائي) في هذا المثال هي الدرجة فإذا رمزنا لها بالرمز X، بما أن المتغير العشوائي هو متغير متقطع، فنحصل على التوزيع الاحتمالي لهذا المجتمع بكتابة القيم المختلفة للمتغير العشوائي وأمام كل قيمة نكتب احتمالها، حيث القيم المختلفة التي يأخذها هذا المتغير العشوائي (الدرجة) هي :

7 ، 6 ، 5 ، 4

ونحسب احتمال كل قيمة من هذه القيم، بقسمة عدد مرات تكرار القيمة على العدد الكلي للقيم، فمثلا عندها $\mathbf{x = 4}$ ، فإن:

$$f(4) = P(X = 4) = \frac{2}{10} = 0.2$$

وهكذا .... فسيكون التوزيع الاحتمالي لهذا المجتمع كما يلي:

جدول (2–4) التوزيع الاحتمالي للمجتمع

x	4	5	6	7	المجموع
$f(x)$	0.2	0.3	0.3	0.2	1

ومن التوزيع الاحتمالي للمجتمع نحسب الوسط الحسابي للمجتمع كما يلي:

$$\mu = \sum_{i=1}^{i=4} x_i\, f(\, x_i)$$

$$= (4)(0.2) + (5)(0.3) + (6)(0.3) + (7)(0.2) = 5.5$$

أي أن الوسط الحسابي للدرجات هو **5.5** درجة

مثال)2-10(:

إذا كان التوزيع الاحتمالي للمتغير العشوائي **X** كما يلي:

x	0	1	2	3	4
$f(x)$	0.01	0.15	0.29	0.35	0.20

أحسب الوسط الحسابي لهذا التوزيع الاحتمالي.

الحل

من التوزيع الاحتمالي نحسب الوسط الحسابي **μ** كما يلي:

$$\mu = \sum_{i=1}^{i=5} x_i\, f(\, x_i)$$

$$= (0)(0.01) + (1)(0.15) + (2)(0.29) + (3)(0.35) + (4)(0.20) = 2.58$$

مثال)2-11(:

أذا ألقينا قطعة نقدية واحدة مرتين، وكان المتغير العشوائي **X** يمثل عدد المرات التي نحصل فيها على وجه، احسب الوسط الحسابي لهذا المتغير .

الحل:

نوجد أولا التوزيع الاحتمالي لهذا المتغير، ثم نحسب منه الوسط الحسابي، ولإيجاد التوزيع الاحتمالي، يجب كتابة فراغ العينة لهذه التجربة العشوائية.

حيث:

S = { HH, HT , TH , TT }

إذن القيم التي يمكن ان يأخذها المتغير العشوائي الذي يمثل عدد المرات التي نحصل فيها على وجه عند إلقاء قطعة نقدية مرتين هي: 0،1،2 .

والتوزيع الاحتمالي لهذا المتغير العشوائي يوضحه الجدول التالي:

x	0	1	2
$f(x)$	1/4	2/4	1/4

ومن التوزيع الاحتمالي نحسب الوسط الحسابي كما يلي:

$$\boldsymbol{\mu = \sum_{i=1}^{i=3} x_i f(x_i)}$$

$$\boldsymbol{= (0)(1/4) + (1)(2/4) + (2)(1/4) = 4/4 = 1}$$

(2-3-2) القيمة المتوقعة:

إذا كان لدينا المتغير العشوائي **X** فالقيمة المتوقعة لهذا المتغير يرمز لها بالرمز E (**X**) وهي عبارة عن الوسط الحسابي لقيم المتغير العشوائي **X** مرجحة باحتمالاتها.

وبالتالي فالقيمة المتوقعة للمتغير العشوائي المنفصل **X**، تحسب كما يلي:

$$\mathrm{E(X)} = \sum_{i=1}^{i=3} x_i f(x_i)$$

وهذه نفسها الصيغة المستخدمة لحساب الوسط الحسابي من التوزيعات الاحتمالية المتقطعة، وبالتالي فالقيمة المتوقعة للمتغير العشوائي المتقطع **X**، هي عبارة عن تسمية اخرى للوسط الحسابي، أي أن:

$$E(\mathbf{X}) = \mu$$

وفي بعض الاحيان نجد أن استخدام مصطلح القيمة المتوقعة يوضح المقصود أكثر من استخدام مصطلح الوسط الحسابي، وذلك كما هو واضح في المثال التالي:

مثال (2-12):

إذا كانت القيمة الاسمية لبوليصة التامين عن حوادث السيارات لشركة تامين معينة **1000** دينار ليبي، وكان قسط التأمين السنوي يساوي **55** دينار ليبي، فإذا كانت نسبة السيارات المتعرضة سنويا للحوادث هي **4%** فما هي القيمة المتوقعة لأرباح هذه الشركة من حامل عقد التأمين في سنة؟

الحل:

إذا اخترنا من حاملي عقود التأمين واحدا عشوائيا وجعلنا المتغير العشوائي **X** يرمز للربح الذي تجنيه الشركة من هذا الشخص في سنة واحدة. فسنجد انه إذا لم يتعرض الشخص لحادث فستربح الشركة في السنة قيمة القسط وهو **55** دينار أي **X = 55** ، أما إذا تعرض الشخص لحادث فستدفع له الشركة مبلغ **1000** دينار، فستكون خسارة الشركة في هذه الحالة (**945−=55+1000−**) أي ان **X=−945** (الاشارة السالبة تعني خسارة)، وحيث أن احتمال وقوع حادث **=40.0** (النسبة هي احتمال)، إذن احتمال عدم وقوع حادث = **0.96**، ويكون التوزيع الاحتمالي لربح الشركة في سنة كما يلي:

x	$f(x)$
55	0.96
−945	0.04

ونلاحظ هنا ان المتغير العشوائي **X** متغير متقطع، إذن فالقيمة المتوقعة لربح هذه الشركة من شخص واحد في سنة يحسب كما يلي:

$$E(X) = \mu = \sum_{i=1}^{i=2} x_i\, f(x_i) = (55)(0.96) + (-945)(0.04) = 15$$

ويعني ذلك ان الشركة تتوقع أن يكون ربحها من حامل عقد التأمين الواحد في سنة **15** دينار. أي متوسط ربح الشركة من الشخص الواحد في سنة **15** دينار.

وبالتالي إذا كان عدد عقود التأمين المبرمة مع هذه الشركة **75** ألف عقد، فستتوقع الشركة ان تكون أرباحها الكلية السنوية تساوي.

75000×15=1125000 دينار.

أي سيكون متوسط الأرباح الكلية السنوية لهذه الشركة مليون ومائة وخمسة وعشرون ألف دينار .

(2-3-3) التباين:

التباين هو مقياس لدرجة تشتت قيم المتغير العشوائي حول وسطها الحسابي ويعرف التباين بانه الوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي.

ويرمز لتباين التوزيع الاحتمالي (لتباين المجتمع) بالرمز σ^2(سيجما تربيع).

ويحسب التباين في حالة التوزيعات الاحتمالية المتقطعة كما يلي:

$$\boxed{\sigma^2 = \sum_{i=1}^{i=r} (x_i - \mu)^2\, f(x_i)}$$

وحيث ان التباين يقيس مدي تشتت قيم التوزيع حول وسطها الحسابي **μ** فعليه إذا كانت قيمة التباين كبيرة فيعني ذلك ان قيم المتغير العشوائي لهذا التوزيع متباعدة عن وسطها **μ**، أما إذا كانت قيمة التباين صغيرة فيعني ذلك أن قيم المتغير العشوائي قريبة ومتمركزة حول وسطها الحسابي **μ**، وبالتالي فعندما يساوي التباين

الصفر فيعني ذلك أنه لا يوجد تشتت، أي أن جميع قيم التوزيع تساوي الوسط الحسابي **μ**، فمثلاً إذا كان لدينا توزيع احتمالي لدرجات طلبة في مادة الإحصاء وعلمت أن الوسط الحسابي لهذا التوزيع **μ=62** وتباين التوزيع $\sigma^2 = 0$، فيعني ذلك أن جميع الطلبة درجاتهم **62**.

يجب الانتباه أن التباين هو مقياس مربع، وبالتالي قيمته ستكون صفر أو قيمة موجبة، **أي أن** $\boldsymbol{\sigma^2 \geq 0}$.

والصيغة السابقة لحساب التباين تسمى صيغة الانحرافات عن الوسط الحسابي، ومنها نستطيع استنتاج صيغة أخرى لغرض تسهيل العمليات الحسابية وخاصة عندما تكون قيمة الوسط الحسابي كسرا، وتسمى صيغة القيم مباشرة، وهي:

$$\boxed{\boldsymbol{\sigma^2 = \left[\sum_{i=1}^{i=r} x_i^2 \ f(x_i)\right] - \mu^2}}$$

وبالطبع تعطي الصيغتان نفس النتيجة تماماً. وبما أن التباين يتعامل مع مربع انحرافات القيم عن وسطها الحسابي، إذن ستكون وحداته هي مربع وحدات القياس الأصلية، وكثيراً ما تكون غير ذات معني، فمثلاً إذا كان المتغير العشوائي المدروس يمثل عدد الأطفال فستكون وحدات التباين طفل تربيع، وإذا كان المتغير يمثل الوقت بالساعات فستكون وحدات التباين ساعة تربيع وكلها ليس لها أي معنى، وهذا يمثل صعوبة في تفسير التباين. وللتخلص من هذه الصعوبة، نُرجع الوحدات إلى أصلها بأخذ الجذر التربيعي للتباين، ويسمى المقياس الجديد بالانحراف المعياري.

(2–3–4) الانحراف المعياري:

الانحراف المعياري هو الجذر التربيعي الموجب للوسط الحسابي لمربعات انحرافات القيم عن وسطها الحسابي، أي الانحراف المعياري هو الجذر التربيعي الموجب للتباين، ويرمز له بالرمز $\boldsymbol{\sigma}$:

$$\boldsymbol{\sigma} = \sqrt{\boldsymbol{\sigma}^2}$$

مثال (2-13):

أحسب التباين والانحراف المعياري للتوزيع الاحتمالي المتقطع التالي:

x	2	4	6	8	10
$f(x)$	0.05	0.15	0.25	0.35	0.20

الحل:

لحساب التباين يجب أولاً حساب الوسط الحسابي $\boldsymbol{\mu}$، حيث:

$$\mu = \sum_{i=1}^{i=5} x_i \ f(x_i) = 2(0.05) + 4(0.15) + 6(0.25) + 8(0.35) + 10(0.20) = 7.0$$

ونحسب التباين باستخدام صيغة الانحرافات عن الوسط الحسابي كما يلي:

$$\sigma^2 = \sum_{i=1}^{i=5} (x_i - \mu)^2 \ f(x_i) = (2-7)^2(0.05) + (4-7)^2(0.15) + (6-7)^2(0.25)$$

$$+ (8-7)^2 \ (0.35) + (10-7)^2 \ (0.20) = 5.0$$

ونستطيع حساب التباين باستخدام صيغة القيم مباشرة كما يلي:

$$\sigma^2 = \sum_{i=1}^{i=5} x_i^2 \ f(xi) - \mu^2 = [(2)^2 \ (0.05) + (4)^2(0.15) + (6)^2(0.25) + (8)^2 \ (0.35) + (10)^2 \ (0.20)] - (7)^2$$

$$= 54.0 - 49.0 = 5.0$$

إذن الانحراف المعياري لهذا المتغير العشوائي يساوي:

$$\sigma = \sqrt{\sigma^2} = \sqrt{5} = 2.236$$

مثال (2-14):

إذا ألقينا قطعة نقدية واحدة مرتين، وكان المتغير العشوائي X يمثل عدد المرات التي نحصل فيها على وجه، أحسب التباين والانحراف المعياري لهذا المتغير.

الحل:

من مثال (2−6)، علمنا أن التوزيع الاحتمالي لهذا المتغير العشوائي كما يلي:

x	0	1	2
$f(x)$	1/4	2/4	1/4

ومن التوزيع الاحتمالي نحسب الوسط الحسابي μ كما يلي:

$$\mu = \sum_{i=1}^{i=3} x_i f(x_i) = (0)(1/4) + (1)(2/4) + (2)(1/4) = 4/4 = 1$$

حساب التباين باستخدام صيغة الانحرافات عن الوسط الحسابي:

$$\sigma^2 = \sum_{i=1}^{i=3} (x_i - \mu)^2 f(x_i)$$

$$= (0-1)^2\left(\frac{1}{4}\right) + (1-1)^2\left(\frac{2}{4}\right) + (2-1)^2\left(\frac{1}{4}\right) = \frac{1}{4} + 0 + \frac{1}{4} = \frac{1}{2}$$

ونستطيع حساب التباين باستخدام صيغة القيم مباشرة كما يلي:

$$\sigma^2 = \left[\sum_{i=1}^{i=3} x_i^2 f(x_i)\right] - \mu^2$$

$$= \left[(0)^2\left(\frac{1}{4}\right) + (1)^2\left(\frac{2}{4}\right) + (2)^2\left(\frac{1}{4}\right)\right] - (1)^2 = \frac{6}{4} - 1 = \frac{1}{2}$$

إذن الانحراف المعياري لهذا المتغير العشوائي يساوي:

$$\sigma = \sqrt{\sigma^2} = \sqrt{\frac{1}{2}} = 0.7071$$

ملخص الفصل الثاني

عرفنا في هذا الفصل المتغير العشوائي بأنه متغير كمي نعتمد في تحديد قيمه على كل نتيجة من النتائج التي يمكن أن نحصل عليها من إجراء تجربة عشوائية، ويوجد نوعان من المتغيرات العشوائية وهما المتغير العشوائي المتقطع (المنفصل) وهو متغير عشوائي قيمه التي يمكن أن يأخذها منفصلة عن بعضها وقابلة للعدّ. والمتغير العشوائي المستمر (المتصل) وهو متغير عشوائي يمكنه أن يأخذ أية قيمة في فترة معينة، أي تكون القيم التي يمكن أن يأخذها متصلة ببعضها أي مستمرة في فترة معينة، وبالتالي فهي غير قابلة للعد.

وإذا كان المتغير العشوائي متقطعـًا، فتوزيعه الاحتمالي يسمى توزيع احتمالي متقطع ونعبر عنه بالدالة$f(x)$ وتسمي دالة كتلة الاحتمال، حيث قيمتها عند قيمة معينة x، هي احتمال أن يأخذ المتغير هذه القيمة أي :$P(X = x) = f(x)$، ويجب أن يتوفر في دالة كتلة الاحتمال الشرطان التاليان:$0 \leq f(x) \leq 1$ ، $\sum f(x) = 1$.

أما إذا كان المتغير العشوائي مستمراً، فتوزيعه الاحتمالي يسمى توزيع احتمالي مستمر ونعبر عنه بالدالة$f(x)$ وتسمي دالة كثافة الاحتمال. واحتمال أن يأخذ المتغير المستمر أية قيمة داخل فترة معينة مساويـًا للمساحة المحصورة بين محور السينات ومنحني دالة كثافة الاحتمال لهذه الفترة. ويجب أن يتوفر في دالة كثافة الاحتمال الشرطان التاليان:$f(x) \geq 0$، والمساحة الكلية المحصورة بين منحنى دالة كثافة الاحتمال ومحور السينات تساوي 1.

ثم تعرضنا لوصف التوزيعات الاحتمالية (المجتمعات)، بحساب بعض معالمها الهامة، وقد حسبنا هذه المعالم عندما يكون توزيع المجتمع توزيعـًا احتماليـًا مقطعـًا، **حيث:**

يُحسب الوسط الحسابي للمجتمع **μ** كما يلي: $\mu = \sum_{i=1}^{i=r} x_i \, f(x_i)$

ويُحسب تباين المجتمع σ^2 بإحدى الصيغتين التاليتين:

$$\sigma^2 = \left[\sum_{i=1}^{i=r} x_i^2 f(x_i)\right] - \mu^2 \quad \text{،} \quad \sigma^2 = \sum_{i=1}^{i=r} (x_i - \mu)^2 \, f(x_i)$$

أما الانحراف المعياري **σ** فهو الجذر الموجب للتباين أي: $\sigma = \sqrt{\sigma^2}$

تمارين (2-2)

1. أحسب الوسط الحسابي والتباين والانحراف المعياري لكل من التوزيعات التالية:

أ-

x	4	5	6	7	8
$f(x)$	0.10	0.15	0.35	0.30	0.10

ب-

x	0	1	2	3	4
$f(x)$	0.14	0.20	0.30	0.35	0.01

جـ-

x	-2	-1	0	1	2	3
$f(x)$	1/12	1/12	2/12	4/12	2/12	2/12

2. عند إلقاء 3 قطع نقدية، واعتبار أن المتغير العشوائي X يمثل عدد الأوجه التي نحصل عليها، أحسب الوسط الحسابي والتباين والانحراف المعياري لهذا المتغير.

3. أحسب الوسط الحسابي والتباين والانحراف المعياري للتوزيع الاحتمالي المتقطع التالي:

$$f(x) = \begin{cases} \frac{1}{7} & x = 0, 1, 2, 3, 4, 5, 6 \\ 0 & otherwise \end{cases}$$

4. إذا كان التوزيع الاحتمالي للمتغير العشوائي X الذي يمثل المبيعات الأسبوعية من السيارات لوكالة معينة كما يلي:

x	0	1	2	3	4
$f(x)$	0.01	0.15	0.29	0.35	0.20

أحسب القيمة المتوقعة لعدد السيارات التي تبيعها هذه الوكالة أسبوعيًا.

5. إذا كان لدينا التوزيع الاحتمالي المتقطع التالي:

x	**0**	**1**	**2**	**3**	**4**
$f(x)$	0.20	k	0.30	0.10	c

وكانت $E(X) = 1.6$ ، فأوجد قيمة K و C.

الفصل الثالث
توزيعات احتمالية هامّة

(3-1) توزيعات احتمالية متقطعة هامة:

لقد درسنا التوزيعات الاحتمالية بصفة عامة، وعرفنا كيفية تحديد صفاتها وذلك باستخدام مقاييس للنزعة المركزية ومقاييس للتشتت. وفي هذا البند سنتعرض لدراسة توزيعين احتماليين من النوع المتقطع، نظراً لأهميتها، والتي تأتي من تطبيقهما بشكل واسع في الدراسات الإحصائية، وهما توزيع ذات الحدين وتوزيع بواسون.

(3-1-1) توزيع ذات الحدين:

إذا كررنا تجربة n من المرات المستقلة حيث يطلق على كل مرة محاولة وكان احتمال النجاح ثابت في جميع المحاولات، فهذه التجربة يطلق عليها تجربة ذات الحدين، أي أن أية تجربة عشوائية تتوفر فيها الشروط التالية تسمى تجربة ذات الحدين:

1) تجربة عشوائية تتكون من n من المحاولات المتماثلة.

2) كل محاولة تصنف نتيجتها إلى نجاح أو فشل، حيث المقصود بالنجاح هو ظهور نتيجة مرغوب فيها، والمقصود بالفشل هو ظهور نتيجة غير مرغوب فيها.

3) جميع المحاولات مستقلة عن بعضها البعض.

4) احتمال النجاح، أي احتمال ظهور النتيجة المرغوب فيها، ثابت من محاولة إلى أخرى.

والمتغير العشوائي X الذي يمثل عدد المحاولات الناجحة يسمى متغير ذات الحدين وتوزيعه الاحتمالي يطلق عليه توزيع ذات الحدين. والقيم التي يمكن أن يأخذها المتغير العشوائي X هي:

0,1,2,3,…………………,n

فعندما يأخذ المتغير X القيمة **0** يعني ذلك أن عدد المحاولات الناجحة يساوي صفر، أي أن كل المحاولات فاشلة، وعندما تكون قيمة المتغير X تساوي 1، يعني ذلك أن محاولة واحدة فقط كانت ناجحة، وهكذا ... وإذا كانت قيمة المتغير X تساوي n يعني ذلك أن كل المحاولات كانت ناجحة.

وبما أن القيم التي يمكن أن يأخذها هذا المتغير العشوائي هي قيم منفصلة وقابلة للعدّ، فعددها يساوي **(n+1)** قيمة، إذن متغير ذات الحدين هو متغير متقطع. وبالتالي توزيعه الاحتمالي هو توزيع احتمالي متقطع، ودالة كتلة الاحتمال لتوزيع ذات الحدين كما يلي:

دالة كتلة الاحتمال لتوزيع ذات الحدين تأخذ الصيغة التالية:

$$\boldsymbol{f(x;n,p)=c_x^n p^x q^{n-x}} \qquad \boldsymbol{x=0,1,2\,\ldots\ldots\ldots\ldots\ldots\ldots n}$$

n, **P** هما المعلمتان اللتان يعتمد عليهما توزيع ذات الحدين. حيث:

n: عدد المحاولات الكلية .

P: احتمال النجاح في المحاولة الواحدة.

q: احتمال الفشل في المحاولة الواحدة، q = (1-p).

x : عدد المحاولات الناجحة.

الوسط الحسابي والتباين لتوزيع ذات الحدين:

إذا كان المتغير العشوائي X يتبع توزيع ذات الحدين، فإن:

الوسط الحسابي: $\boldsymbol{\mu = np}$

التبايـــــــــــن: $\boldsymbol{\sigma^2}$**=n p q = n p(1-p)**

مثال)1-3(:

إذا كان المتغير العشوائي X يتبع توزيع ذات الحدين بمعلمتين:

P=1/3، **n =5**

أ) أحسب احتمال أن يساوي المتغير القيمـــة **1** .

ب) أحسب الوسط الحسابي والتباين لهذا التوزيع.

الحل:

أ. بما أن المتغير يتبع توزيع ذات الحدين، إذن:

$$f(x;n,p) = C_x^n P^x q^{n-x} \quad x = 0,1,2 \ldots\ldots\ldots\ldots..n$$

وبما أن **n =5** ،**P=1/3** ، إذن في هذه الحالة دالة كتلة الاحتمال كما يلي:

$$f(x;5,1/3) = C_x^5(\frac{1}{3})^x(\frac{2}{3})^{5-x} \qquad x = 0,1,2,3,4,5$$

وبالتعويض في صيغة دالة كتلة الاحتمال بالقيمة $x = 1$ نحصل على الاحتمال المطلوب وذلك كما يلي:

$$P(X = 1) = f(1) = C_1^5\ (\frac{1}{3})^1(\frac{2}{3})^4 = 80/243 = 0.3292$$

ب. الوسط الحسابي: $\mu = np = (5)(1/3) = 5/3 = 1\frac{2}{3}$

التباين: $\sigma^2 =$ npq = (5)(1/3)(2/3) = $\frac{10}{9}$ = 1 $\frac{1}{9}$

مثال)2-3(:

أشترك **7** طلبة في امتحان في مادة الرياضيات، فإذا كان احتمال النجاح في هذا الامتحان **0.60**، أحسب:

أ) احتمال أن ينجح 4 طلبـــة.

ب) احتمال أن ينجح كل الطلبة.

الحل:

في هذا المثال نلاحظ أن التجربة العشوائية المذكورة تتوفر فيها شروط توزيع ذات الحدين حيث ان عدد المحاولات الكلية = العدد الكلي للطلبة المشتركين في الامتحان، أي أن : **n = 7**

واحتمال النجاح **P = 0.60**

بما أن دالة كتلة الاحتمال لتوزيع ذات الحدين في هذا المثال كما يلي:

$$\boldsymbol{f(x; 7, 0.6) = C_x^7 (0.6)^x (0.4)^{7-x} \qquad x = 0, 1, 2, \dots \dots \dots . . 7}$$

أ) نحصل على احتمال أن ينجح 4 طلبة بالتعويض في صيغة دالة كتلة الاحتمال بالقيمة، وذلك كما يلي:

$$\boldsymbol{P(X = 4) = f(4) = C_4^7 (0.6)^4 (1 - 0.6)^{7-4}}$$

$$\boldsymbol{= \frac{7!}{4!(7-4)!} (0.6\,)^4 \ (0.4)^3 \ = 0.2903}$$

ب) نحصل على احتمال ان ينجح كل الطلبة بالتعويض في صيغة دالة كتلة الاحتمال بالقيمة $\boldsymbol{x = 7}$ ، وذلك كما يلي:

$$\boldsymbol{P(x = 7) = f(7) = C_7^7\, (0.6)^7 (1 - 0.6)^{7-7}}$$

$$\boldsymbol{= \frac{7!}{7!(7-7)!}\ (0.6)^7\ (0.4)^0\ = (0.6)^7\ = 0.028}$$

مثال)3-3(:

إذا كان 0.10 من الإنتاج الكلي في مصنع معين إنتاجا تالفًا، فإذا سحبنا عشوائيًا من هذا الإنتاج 6 وحدات، فما احتمال أن يكون عدد الوحدات التالفة أقل من3 وحدات؟

الحل:

بفرض أن المتغير العشوائي X يمثل عدد الوحدات التالفة، فسيكون X متغير عشوائي يتبع توزيع ذات الحدين بالمعلمتين n=6 ، P=0.10 ، والاحتمال المطلوب: $P(X<3)$ حيث:

$$P(X<3)=P(X=0)+P(X=1)+P(X=2)$$

وبالتعويض في دالة كتلة الاحتمال لتوزيع ذات الحدين، نحصل على الاحتمالات التالية:

$$P(X=0)=f(0)=C_0^6(0.10)^0(1-0.10)^{6-0}$$
$$=\frac{6!}{0!\,(6-0)!}(1)(0.90)^6=0.5314$$

$$P(X=1)=f(1)=C_1^6(0.10)^1(1-0.10)^{6-1}$$
$$=\frac{6!}{1!\,(6-1)!}(1.10)(0.90)^5=0.3542$$

$$P(X=2)=f(2)=C_2^6(0.10)^2(1-0.10)^{6-2}$$
$$=\frac{6!}{2!\,(6-2)!}(1.10)^2(0.90)^4=0.0984$$

إذن الاحتمال المطلوب هو:

$$P(X<3)=P(X=0)+P(X=1)+P(X=2)$$
$$=0.5314+0.3542+0.0984=0.9840$$

مثال)4-3(:

إذا ألقينا مكعب نرد **4** مرات، فما احتمال ظهور العدد **3** مرتين أو أكثر؟

الحل:

هنا تعتبر الرمية أي المحاولة ناجحة إذا ظهر العدد **3**، فاحتمال الحصول على العدد **3** مرتين أو أكثر المقصود به أن يكون عدد المحاولات الناجحة **2** أو أكثر، أي يكون المتغير العشوائيX الذي يمثل عدد المحاولات الناجحة يساوي **2** أو أكثر، بما أن هذا المتغير يتبع توزيع ذات الحدين بمعلمتين:

n=4 (عدد المحاولات الكلية).

P=1/6 (احتمال النجاح = احتمال ظهور العدد 3 في محاولة واحدة).

إذن دالة الاحتمال لهذا المتغير العشوائي X:

$$f(x;4,1/6)=C_x^4\left(\frac{1}{6}\right)^x\left(\frac{5}{6}\right)^{4-x}\qquad x=0,1,2,3,4$$

والاحتمال المطلوب يحسب كما يلـــــي:

$$P(X\geq 2)=f(2)+f(3)+f(4)$$

حيث:

$$P(X=2)=f(2)=C_2^4\left(\frac{1}{6}\right)^2\left(\frac{5}{6}\right)^2=0.1157$$

$$P(X=3)=f(3)=C_3^4\left(\frac{1}{6}\right)^3\left(\frac{5}{6}\right)=0.0154$$

$$P(X=4)=f(4)=C_4^4\left(\frac{1}{6}\right)^4\left(\frac{5}{6}\right)^0=0.0008$$

إذن الاحتمال المطلوب:

$$P(X\geq 2)=f(2)+f(3)+f(4)=0.1157+0.0154+0.0008=0.1319$$

ملاحظة:

بما ان$\sum f(x)=1$، إذن نستطيع حساب الاحتمال المطلوب في هذا المثال كما يلي:

$$P(X \geq 2) = 1 - P(X < 2)$$

ويكون من السهل الحل باستخدام هذه الطريقة، عندما يكون العدد الكلي للمحاولات n كبيراً، فمثلاً إذا كانت n=10، **فإن** $P(X \geq 2)$ يحسب كما يلي:

$$P(X \geq 2) = 1 - P(X < 2) = 1 - [f(0) + f(1)]$$

أسهل من حسابه كما يلي:

$$P(X \geq 2) = f(2) + f(3) + f(4) + f(5) + f(6) + f(7) + f(8) + f(9) + f(10)$$

(3-1-2) توزيع بواسون:

التجارب العشوائية التي تعطينا نتائجها قيمة المتغير العشوائي X الذي يمثل العدد الكلي لمرات الحصول على حدث ما في فترة زمنية معينة، قد تكون هذه الفترة ثانية أو دقيقة أو ساعة أو يومـاً أو ... الخ. أو في منطقة محددة، قد تكون هذه المنطقة صفحة من كتاب أو متراً مربعـاً من مساحة ... الخ، بحيث يكون ظهور الحدث في هذه الفترة أو المنطقة عشوائيـاً، هذه التجارب يطلق عليها تجارب بواسون، نسبة للعالم الفرنسي سيمون بواسون، والمتغير العشوائي X يطلق عليه متغير بواسون ـ وتوزيعه الاحتمالي يسمى توزيع بواسون.

ومن الأمثلة على تجربة بواسون، عدد الزبائن الذين يدخلون محل تجاري معين في الساعة، عدد السيارات التي تمر من أمام أحد الفنادق في الساعة، عدد حوادث السيارات في مدينة ما شهريـاً، عدد الأخطاء المطبعية في صفحة من كتاب، عدد الفئران في هكتار من منطقة زراعية، وبما أننا نحصل على قيم متغير بواسون عن طريق العد، إذن متغير بواسون هو متغير عشوائي متقطع، وبالتالي سيكون له دالة كتلة احتمال، وهي كما يلي:

دالة الكتلة الاحتمالية لتوزيع بواسون تأخذ الصيغة التالية:

$$f(x;\lambda)=\frac{e^{-\lambda}\lambda^{x}}{x!}\qquad x=0,1,2,\ldots\ldots$$

يعتمد التوزيع على معلمة واحدة وهي λ ، حيث:

λ : هي معدل وقوع الحدث في وحدة زمنية أو مكانية واحدة وهي تساوي الوسط الحسابي لتوزيع بواسون.

x: العدد الكلي لمرات ظهور ظاهرة ما في فترة زمنية معينة.

$2.71828\cong e$

الوسط الحسابي والتباين لتوزيع بواسون:

إذا كان المتغير العشوائي X يتبع توزيع بواسون، فإن:

وسطه الحسابي: $\mu=\lambda$

وتباينــــــــــه : $\sigma^2=\lambda$

أي أن الوسط الحسابي للتوزيع = تباينه = λ

مثال (3-5):

إذا علمت أن الزبائن يدخلون محل للملابس بمعدل **6** زبائن في الساعة، فأحسب ما يلي:

أ) احتمال أن يدخل المحل **8** زبائن خلال الساعــــة القادمــــــة.

ب) احتمال أن يدخل المحل ما بين **3** و**7** زبائن خلال الساعة القادمة.

جـ) احتمال أن يدخل المحل أكثر من **4** زبائن خلال الساعة القادمـة.

الحل:

المتغير العشوائي الذي يمثل عدد الزبائن الذين يدخلون المحل في الساعة الواحد هو متغير بواسون، **حيث:** $\boldsymbol{\lambda=\mu=6}$

إذن:

أ) $\boldsymbol{P(X=8)=f(8)=\frac{e^{-6}(6)^8}{8!}=0.1033}$

ب) $\boldsymbol{P(3<X<7)=f(4)+f(5)+f(6)}$

$$= \frac{e^{-6}(6)^4}{4!} + \frac{e^{-6}(6)^5}{5!} + \frac{e^{-6}(6)^6}{6!}$$

$$= 0.1339 + 0.1606 + 0.1606 = 0.4551$$

ج) بما أن توزيع بواسون كأى توزيع متقطع، يجب أن يكون $\sum f(x) = 1$ ،إذن

نستطيع حساب الاحتمال المطلوب كما يلي:

$$P(X > 4) = 1 - [f(0) + f(1) + f(2) + f(3) + f(4)]$$

$$= 1 - \left[\frac{e^{-6}(6)^0}{0!} + \frac{e^{-6}(6)}{1!} + \frac{e^{-6}(6)^2}{2!} + \frac{e^{-6}(6)^3}{3!} + \frac{e^{-6}(6)^4}{4!}\right]$$

$$=1-[0.0025 + 0.0149 + 0.0446 + 0.0892 + 0.1339]=0.7149$$

مثال (3-6) :

إذا علمت أن بدَّالة امعة ما تستقبل في فترة دوام الموظفين في المتوسط

مكالمتين في الدقيقة الواحدة، فأحسب ما يلي:

أ) احتمال أن تستقبل هذه البدّالة مكالمة واحدة في الدقيقة القادمــة.

ب) احتمال ان تستقبل هذه البدّالة مكالمتين أو أكثر في الدقيقة القادمة.

الحل:

المتغير العشوائي الذي يمثل عدد المكالمات التي تستقبلها البدّالة في الدقيقة

الواحدة هو متغير بواسون، حيث:

$$\lambda = \mu = 2$$

أ.

$$P(X = 1) = f(1) = \frac{e^{-2}(2)}{1!} = 0.2707$$

ب.

$$P(X \geq 2) = 1 - [f(0) + f(1)]$$

$$= 1 - \left[\frac{e^{-2}(2)^0}{0!} + \frac{e^{-2}(2)}{1!}\right]$$

$$= 1 - [0.1353 + 0.2707] = 0.5940$$

تمارين (3–1)

1. إذا إلقينا قطعة نقود **5** مرات، فأحسب احتمال الحصول على وجه أكثر من مرتين.

2. إذا إلقينا زهرة نرد **8** مرات، فأحسب ما يلـــــي:

أ- احتمال الحصول على العدد **6** أربع مـــــــــرات.

ب- احتمال الحصول على عدد أكبر من **4** ثلاث مرات.

جـ- احتمال الحصول على العدد **5** أكثر من مرتيــــن.

د- احتمال الحصول على عدد زوجي مرتين أو أقـــل.

3. إذا علمت أن احتمال أن تكون وحدة تالفة في إنتاج مصنع ما يساوي **0.04**، فإذا استلمنا طلبية من إنتاج هذا المصنع تحتوي على **7** وحدات، فأحسب احتمال أن تحتوي هذه الطلبية على وحدة فقط تالفة.

4. في التمرين السابق، إذا كانت الطلبية تحتوي على **100** وحدة، فأحسب الوسط الحسابي وتباين عدد الوحدات التالفة في هذه الطلبية.

5. أشترك **6** طلبة في امتحان مادة الرياضة، فإذا كان احتمال النجاح في هذا الامتحان **0.72** أحسب ما يلي:

أ) احتمال أن ينجح **4** طلبـــــــة.
ب) احتمال أن ينجح كل الطلبــــة.
جـ) احتمال أن لا ينجح أحــــــد .
د) احتمال أن ينجح أكثر من **5** طلبة.

6. إذا كان معدل الوحدات التي بها عيوب في إنتاج آلة في الساعة الواحدة يساوي **2** ، فأحسب ما يلي:

أ. احتمال أن تنتج هذه الآلة **4** وحدات بها عيوب في الساعة القادمـــة.

ب. احتمال أن لا تنتج هذه الآلة أية وحدة بها عيوب في الساعة القادمة.

7. إذا علمت أن الوسط الحسابي للأخطاء الإملائية التي يرتكبها طالب في السنة الخامسة الابتدائية يساوي **5** أخطاء في الصفحة الواحدة، فإذا طلبنا من طالب في السنة الخامسة أن يكتب صفحة واحدة، فما احتمال أن يكون عدد الأخطاء التي يرتكبها **2** أو أقل؟

8. إذا علمت أن تباين عدد الزبائن الذين يدخلون محل مواد غذائية في الدقيقة الواحدة 3 زبائن، فأحسب ما يلي:

أ. احتمال أن يدخل المحل **4** زبائن في الدقيقة القادمة.

ب. احتمال أن يدخل المحل ما بين **2** و**5** زبائن في الدقيقة القادمة.

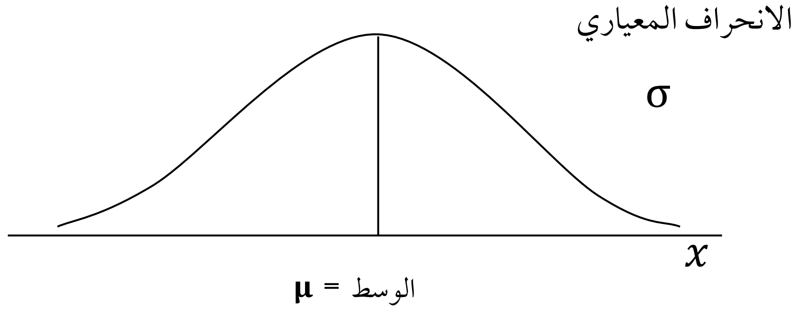
جـ. احتمال أن يدخل المحل أكثر من **3** زبائن في الدقيقة القادمة.

(3–2) توزيعات احتمالية مستمرة هامّة:

يوجد عدد لا نهائي من التوزيعات المستمرة، ولكن أهمها هو التوزيع الطبيعي وبعض التوزيعات المستمرة الأخرى المتعلقة به، مثل توزيع **t** ، والتي سنقوم بدراستها فيما يلي :

(3–2–1) التوزيع الطبيعي :

التوزيع الطبيعي من أهم التوزيعات الاحتمالية المستمرة واكثرها استعمالا في التطبيقات الإحصائية، وقد سمى بالتوزيع الطبيعي، لأن الكثير من الظواهر المشاهدة في العلوم الطبيعية تتبع هذا التوزيع أو يستعمل كتقريب لتوزيعاتها، ومن هذه الظواهر (المتغيرات)، الأطوال، الأوزان، مستوى الذكاء .... الخ. والمنحني الذي يمثل دالة كثافة الاحتمال لهذا التوزيع الاحتمالي المستمر يسمى المنحني الطبيعي، وهو منحني ناقوسي الشكل، أي وحيد المنوال، ويمتد طرفاه إلى ما لانهاية ومتماثل حول قيمة الوسط الحسابي، وذلك كما هو موضح في **شكل (3–1)**.



شكل (3–1)

يطلق على المتغير العشوائي المستمر X الذي يتبع هذا التوزيع، المتغير العشوائي الطبيعي. وتعتمد دالة كثافة الاحتمال لهذا التوزيع على معلمتين، هما الوسط الحسابي للتوزيع μ وتباين التوزيع σ^2.

ودالة كثافة الاحتمال للتوزيع الطبيعي تأخذ الصورة التالية:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

μ ، σ^2 هما المعلمتان اللتان يعتمد عليهما التوزيع الطبيعي ، حيث :

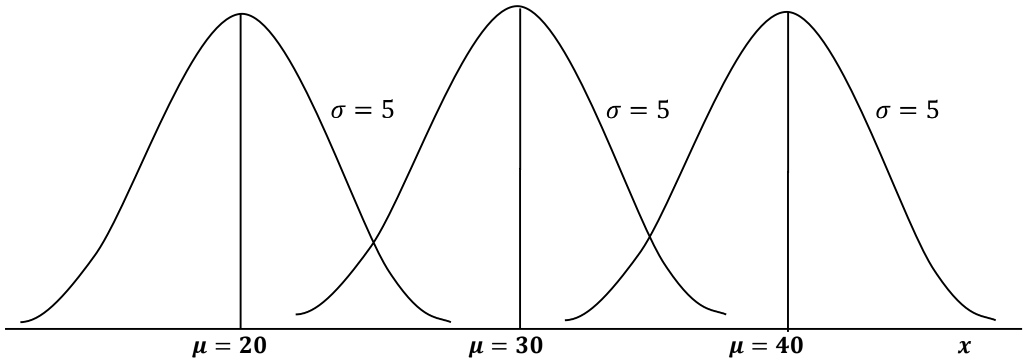
μ : الوسط الحسابي للتوزيع الطبيعي.

σ^2 : تباين التوزيع الطبيعي.

$$e = 2.71828 \ldots \ldots \qquad \pi = 3.14159 \ldots \ldots$$

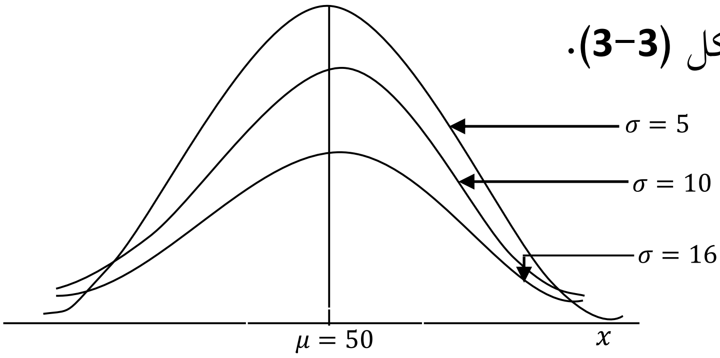
وتوجد عائلة من التوزيعات الطبيعية، تختلف باختلاف الوسط الحسابي والتباين حيث الوسط الحسابي يحدد لنا مركز التوزيع، وبما أن التوزيع الطبيعي هو توزيع متماثل فسنجد أن **الوسط الحسابي** = **الوسيط** = **المنوال**، وبالتالي ستكون قيمة الوسط الحسابي هي قيمة X التي تحت موضع قمة المنحني الطبيعي.

فإذا كان لدينا توزيعان مختلفان في قيمة الوسط الحسابي ومتساويان في التباين فيكون المنحنيان متماثلين تماماً والاختلاف بينهما هو موقع كل منهما، فالتوزيع الذي وسطه الحسابي أكبر يكون موقعه على يمين المنحني الآخر، وذلك كما هو واضح في شكل **(3−2)**.



شكل (3−2)

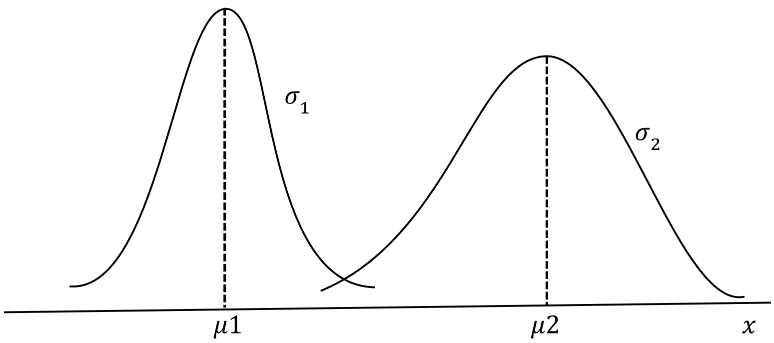
أما إذا كان المنحنيان متساويين في الوسط الحسابي ومختلفين في التباين، فسيكون المنحنيان في نفس الموقع ولكن المنحني الذي يمثل التوزيع الأكثر تبايناً ستكون قمته منخفضة ومتسعة أكثر من المنحني الذي يمثل التوزيع الآخر، وذلك كما هو واضح في شكل **(3−3)**.



شكل (3−3)

وبالطبع قد يختلف التوزيعان في الوسط الحسابي والتباين، وذلك كما هو واضح في شكل **(3−4)**.

$$\boldsymbol{\mu_1 < \mu_2} \quad , \quad \boldsymbol{\sigma_1 < \sigma_2}$$



شكل (3−4)

(3−2−2) خواص التوزيع الطبيعي:

1. المنحني الذي يمثل دالة كثافة الاحتمال للتوزيع الطبيعي، منحنى متصل لجميع قيم x من ∞ − إلى ∞ ويأخذ شكلاً ناقوسياً وحيد المنوال ومتماثلا حول قيمة الوسط الحسابي وهي القيمة المقابلة لقمة المنحني، وذلك كما هو واضح في شكل

(3−4)، وحيث أن المساحة الكلية المحصورة بين منحني أية دالة كثافة احتمال والمحور الأفقي = **1**، إذن فالمساحة على يمين الوسط الحسابي = المساحة التي على يساره = **0.50** .

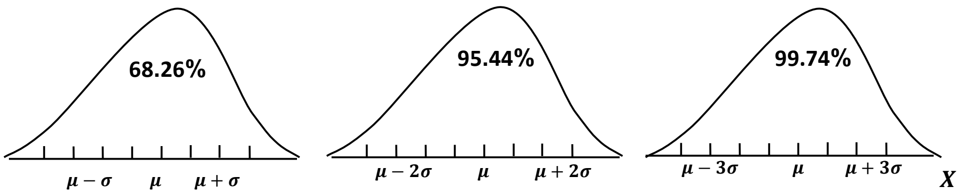
2. الوسط الحسابي للتوزيع = المنوال = الوسيط.
3. يمتد طرفا المنحني إلى ما لا نهاية ويزداد اقترابهما من المحور الأفقي كلما بعدت النقطة عن الوسط الحسابي ولكن لا يلتقيان به.
4. بما أن التوزيع الطبيعي توزيع متماثل، إذن أي معامل للالتواء يساوي **0**.
5. المعامل العزمي للتفرطح للتوزيع الطبيعي = **3** ، أي أن التوزيع الطبيعي هو توزيع معتدل، ولذلك يسمى أحياناً بالتوزيع المعتدل.
6. المساحة تحت المنحني الطبيعي والمحصورة بين $\boldsymbol{\mu - \sigma}$، $\boldsymbol{\mu + \sigma}$ تمثل **68.26%** من المساحة الكلية ، والمساحة بين $\boldsymbol{\mu - 2\sigma, \mu + 2\sigma}$ تمثل **95.44%** والمساحة بين $\boldsymbol{\mu - 3\sigma, \mu + 3\sigma}$ تمثل **99.74%**، وذلك كما هو واضح في شكل **(3−5)**.

وبما أن المساحات تحت دالة كثافة الاحتمال تمثل احتمالات، إذن:

$$\boldsymbol{P(\mu - \sigma < X < \mu + \sigma) = 0.6826}$$

$$\boldsymbol{P(\mu - 2\sigma < X < \mu + 2\sigma) = 0.9544}$$

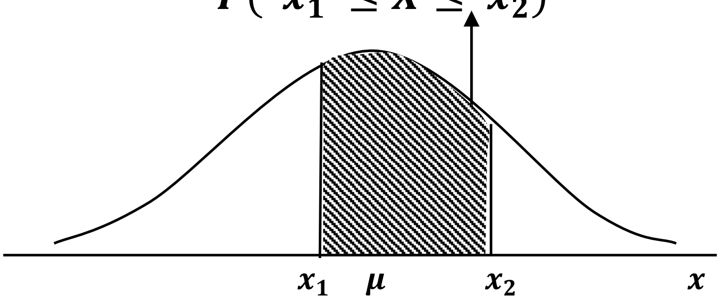
$$\boldsymbol{P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.9974}$$



شكل (3−5)

ونستطيع حساب احتمال أن يأخذ المتغير العشوائي الطبيعي **X** أية قيمة بين قيمتين $\boldsymbol{x_1}$، $\boldsymbol{x_2}$ بحساب المساحة المحصورة بين المنحنى ومحور السينات

والواقعة بين $X = x_1$, $X = x_2$ وذلك بتكامل دالة كثافة الاحتمالات، وذلك كما هو موضح في شكل **(3–6)**.



شكل (3–6)

ولكن هذه الدالة ليس من السهل تكاملها، ولتسهيل حساب الاحتمال بالنسبة لأي توزيع طبيعي نقوم بتحويل المتغير العشوائي الطبيعي **X** إلى متغير عشوائي طبيعي معياري نرمز له بالرمز **Z** ، ويطلق على توزيع المتغير العشوائي **Z** التوزيع الطبيعي المعياري.

(3–2–3) التوزيع الطبيعي المعياري:

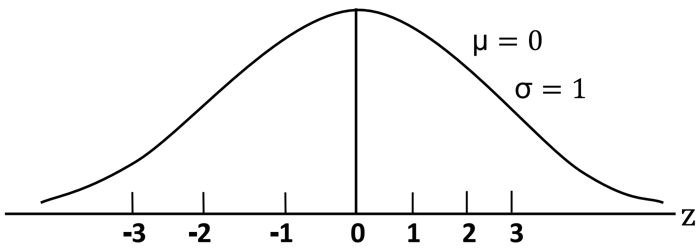
هو توزيع طبيعي وسطه الحسابي $\mu = 0$ وتباينه $\sigma^2 = 1$.

ونقوم بتحويل المتغير العشوائي الطبيعي **x** إلى المتغير العشوائي الطبيعي المعياري **z**، كما يلي:

$$Z = \frac{X-\mu_x}{\sigma_x}$$

بالتعويض في دالة كثافة الاحتمال الخاصة بأي متغير عشوائي طبيعي **X** ، عن الوسط الحسابي **μ** بصفر وعن التباين σ^2 بالقيمة واحد نحصل على دالة كثافة الاحتمال للمتغير العشوائي الطبيعي المعياري **Z** الموضحة في شكل **(3–7)** والتي تأخذ الصورة التالية:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad -\infty < z < \infty$$



شكل (3−7)

والمساحات تحت المنحني الطبيعي المعياري من السهل الحصول عليها بتكامل دالة كثافة الاحتمال $f(Z)$، وقد حُسبت هذه المساحات وعرضت في جدول ، وذلك مثل جدول رقم (م . 1) في ملحق الجداول الإحصائية.

واحتمال أن يأخذ المتغير العشوائي الطبيعي **Z** أية قيمة في الفترة (z_1, z_2) يساوي المساحة المحصورة بين المنحنى الطبيعي المعياري والمحور الأفقي وبين المستقيمين $Z = z_1$ ، $Z = z_2$. ونستخدم جدول **(م.1)** لحساب هذه المساحة أي حساب $P(z_1 < Z < z_2)$.

والبيانات داخل هذا الجدول تمثل المساحات المحصورة بين المنحنى الطبيعي المعياري والمحور الأفقي وبين المستقيم **Z = 0** ، والمستقيم المرسوم عند القيمة المعطاة للمتغير **Z** ، وقيم هذا المتغير يمثلها العمود الأول والسطر الأول في الجدول فالعمود الأول يمثل العدد الصحيح والعدد العشري الأول للمتغير **Z** ، أما السطر الأول فيحدد العدد العشري الثاني لهذا المتغير، وإذا كانت قيمة المتغير **Z** ، تحتوي على أكثر من رقمين عشريين ، يجب تقريبها إلى رقمين عشريين أولا ثم نستخدم هذا الجدول .

واعتمادا على خاصية التماثل التي يتمتع بها المنحني الطبيعي المعياري، نستطيع حساب أي مساحة مرغوب فيها، بين المنحني والمحور الأفقي، وذلك كما سنبينها من خلال الأمثلة التالية :

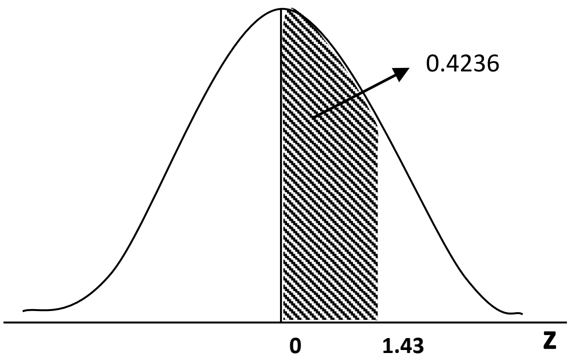
مثال (3-7) :

إذا كان المتغير العشوائي **Z** يتبع التوزيع الطبيعي المعياري، فأحسب الاحتمالات التالية:

أ.	$P(0 \le Z \le 1.43)$	د.	$P(-0.57 \le Z \le 1.12)$
ب.	$P(Z \le 1.43)$	هـ	$P(1.12 \le Z \le 1.41)$
جـ.	$P(-1.43 \le Z \le 1.43)$	و.	$P(Z \ge 1.28)$

الحل:

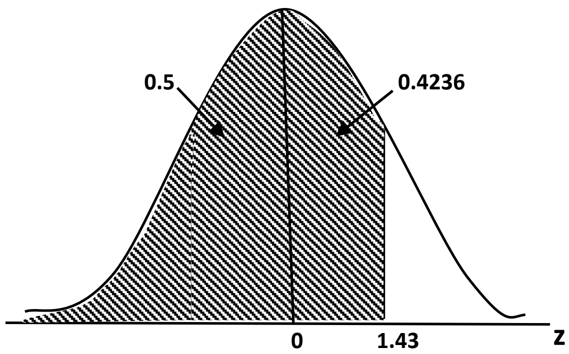
أ- الاحتمال المطلوب $P(0 \le Z \le 1.43)$ يساوي المساحة بين المنحني الطبيعي المعياري ومحور السينات، والمحصورة بين المستقيمين **Z =0** و**Z =1.43** وذلك كما هو واضح من شكل **(3-8)**، ومن جدول (م.**1**)، نجد أن هذه المساحة مساوية للقيمة **0.4236** . أي أن: $P(0 \le Z \le 1.43) = 0.4236$



شكل (3-8)

ب- الاحتمال المطلوب $P(Z \le 1.43)$ يساوي المساحة بين المنحنى الطبيعي المعياري ومحور السينات والتي على يسار القيمة **1.43**، أي المساحة من$-\infty$ إلى **1.43** ، وهي تساوي المساحة من $-\infty$ إلى **0** مضافاً إليها المساحة **0** إلى **1.43**، وذلك كما هو موضح في الشكل **(3-9)**، وبما أن المساحة من $-\infty$ إلى **0** تساوي

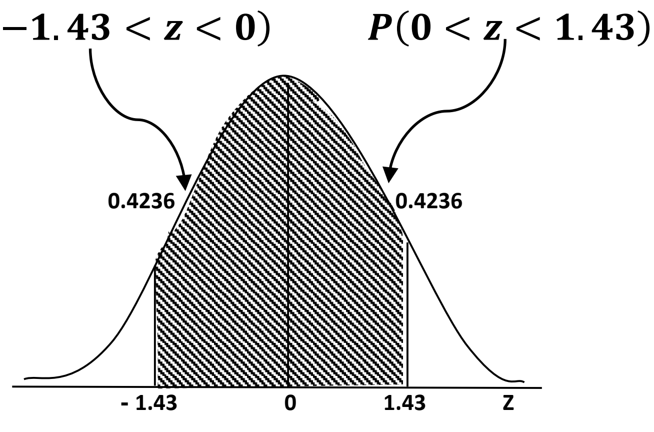
0.50 والمساحة من **0** إلى **1.43** تساوي **0.4236** من جدول (م.**1**)، إذن الاحتمال المطلوب يساوي: $\boldsymbol{P(Z \leq 1.43) = 0.50 + 0.4236 = 0.9236}$



شكل (3−9)

جـ− الاحتمال المطلوب $\boldsymbol{P(-1.43 \leq Z \leq 1.43)}$ يساوي المساحة بين المنحنى الطبيعي المعياري ومحور السينات والمحصورة بين المستقيمين $\boldsymbol{Z = -1.43}$ و$\boldsymbol{Z = 1.43}$، وبما أن المساحة من **0** إلى **1.43** تساوي المساحة من **1.43−** إلى **0**، كما هو واضح من شكل (**3−10**) ، وذلك لأن المنحنى الطبيعي المعياري متماثل ، إذن الاحتمال المطلوب :

$$\boldsymbol{P(-1.43 \leq Z \leq 1.43) = 0.4236 + 0.4236 = 0.8472}$$



شكل (3−10)

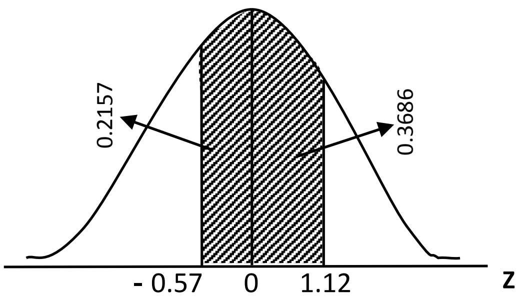
د- الاحتمال المطلوب$P(-0.57 \le Z \le 1.12)$ يساوي المساحة بين المنحنى الطبيعي المعياري ومحور السينات والمحصورة بين المستقيمين **Z = - 0.57** و **Z=1.12**، وهي تساوي المساحة من **0.57-** إلى **0** مضافاً إليها المساحة من **0** إلى **1.12**، وذلك كما هو موضح في الشكل **(3-11)**، وبما ان المساحة من **0.57-** إلى **0** تساوي المساحة من **0** إلى **0.57** لأن المنحنى متماثل ، فمن جدول **(م.1)** نجد أن:

المساحة من **0** إلى **0.57** = **0.2157** .

المساحة من **0** إلى **1.12** = **0.3686** .

إذن الاحتمال المطلوب:

$$P(-0.57 \le Z \le 1.12) = 0.2157 + 0.3686 = 0.5843$$



شكل (3-11)

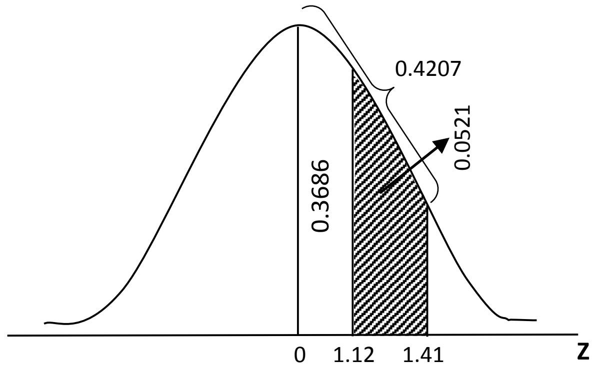
هـ -الاحتمال المطلوب $P(1.12 \le Z \le 1.41)$ يساوي المساحة بين المنحنى الطبيعي المعياري ومحور السينات والمحصورة بين المستقيمين **Z=1.12** و **Z=1.41** وهي تساوي المساحة من **0** إلى **1.41** مطروحاً منها المساحة من **0** إلى **1.12**، وذلك كما هو موضح في الشكل **(3-12)**، وبما أن:

المساحة من **0** إلى **1.41** = **0.4207**

المساحة من **0** إلى **1.12** = **0.3686** .

إذن الاحتمال المطلوب:

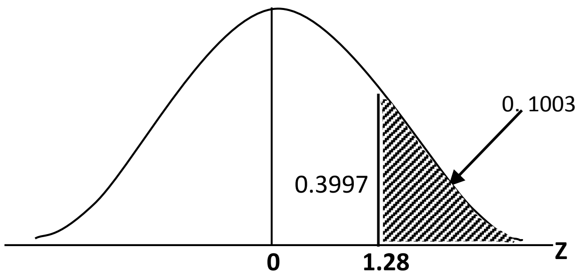
$$P(1.12 \leq Z \leq 1.41) = 0.4207 - 0.3686 = 0.0521$$



شكل (3−12)

و-الاحتمال المطلوب $P(Z \geq 1.28)$ يساوي المساحة بين المنحنى الطبيعي الميعاري ومحور السينات والتي على يمين القيمة **1.28**، أي المساحة من **1.28** إلى ∞ وبما أن المساحة على يمين **0** تساوي **0.50**، إذن المساحة التي تمثل الاحتمال المطلوب ، تساوي **0.50** مطروحاً منه المساحة من **0** إلى **1.28**، وذلك كما هو موضح في الشكل **(3−13)** ، ومن جدول **Z** نجد أن ، المساحة من **0** إلى **1.28** = **0.3997** ، إذن الاحتمال المطلوب يساوي:

$$P(Z \geq 1.28) = 0.50 - 0.3997 = 0.1003$$



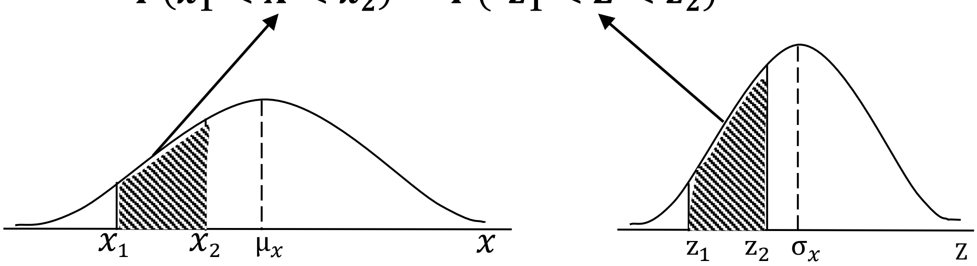
شكل (3−13)

وباستخدام التوزيع الطبيعي المعياري نستطيع تحديد أي احتمال لأي متغير عشوائي طبيعي **X** وذلك لأن المساحة المحصورة تحت دالة كثافة الاحتمال للمتغير X والمحور الأفقي وبين القيمتين $X = x_1$,$X = x_2$ هي نفسها المساحة المحصورة تحت دالة كثافة الاحتمال للمتغير Z وبين القيمتين $Z = z_1$،$Z = z_2$ حيث z_1 هي القيمة المعيارية للقيمة x_1، والقيمة z_2 هي القيمة المعيارية للقيمة x_2 ، أي أن:

$$z_2 = \frac{x_2 - \mu_x}{\sigma_x} \qquad\qquad z_1 = \frac{x_1 - \mu_x}{\sigma_x}$$

وذلك كما هو موضح في شكل **(3−14)**

$$P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$



شكل (3−14)

المساحة المظلة تحت المنحني الطبيعي الأصلي = المساحة المظلة تحت المنحنى المعياري.

$$P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$

حيث :

$$z_1 = \frac{x_1 - \mu_x}{\sigma_x}$$

$$z_2 = \frac{x_2 - \mu_x}{\sigma_x}$$

مثال)3-8(:

إذا كانت أوزان وحدات منتجة من سلعة ما، تتبع توزيعـًا طبيعيـًا بوسط حسابي قدره **65** جرام وتباين قدره **25**، فأحسب احتمال أن يكون وزن وحدة مختارة عشوائيـًا من هذه السلعة:

1. محصور بين **58.75** جرام و **67.50** جرام.

2. أقل من **61.25** جرام.

3. أكثر من **71.25** جرام.

الحل:

1- المطلوب احتمال أن يكون الوزن محصوراً بين **58.75** و **67.50** جرام، أي

الاحتمال المطلوب هو: $P(58.75 < X < 67.50) = ?$

بما أن: $P(58.75 < X < 67.50) = P(z_1 < Z < z_2)$

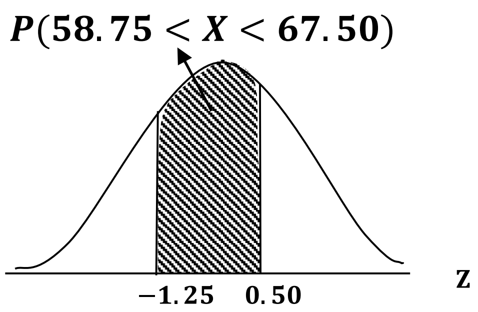
حيث:

$$z_1 = \frac{58.75 - 65}{5} = -1.25$$

$$z_2 = \frac{67.50 - 65}{5} = 0.50$$

$$\therefore \quad P(58.75 < X < 67.50) = P(-1.25 < Z < 0.50)$$

أي أن الاحتمال المطلوب يساوي احتمال أن يقع المتغير بين القيمتين، **0.50**، **1.25**-، وهو يساوي المساحة المحصورة بين هاتين القيمتين، وباستخدام جدول، (م.1)وكما هو واضح في شكل **(3−15)**، نجد أن:



شكل (3−15)

$$P(58.75 < X < 67.50) = P(-1.25 < Z < 0.5)$$
$$= 0.3944 + 0.1915 = 0.5859$$

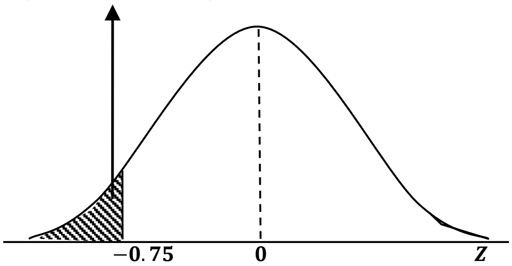
-2 المطلوب احتمال أن يكون وزن الوحدة أقل من **61.25** جرام أي:

$$\mathbf{P(X < 61.25) = ?}$$

نحسب القيمة المعيارية **Z** المقابلة للقيمة **61.25** حيث:

$$\mathbf{Z} = \frac{x - \mu_x}{\sigma_x} = \frac{\mathbf{61.25 - 65}}{\mathbf{5}} = \mathbf{-0.75}$$

إذن: $\mathbf{P(X < 61.25) = P(Z < -0.75) = 0.50 - 0.2734 = 0.2266}$



شكل (3−16)

3− المطلوب هو احتمال أن يكون وزن الوحدة أكثر من **71.25** جرام، أي:

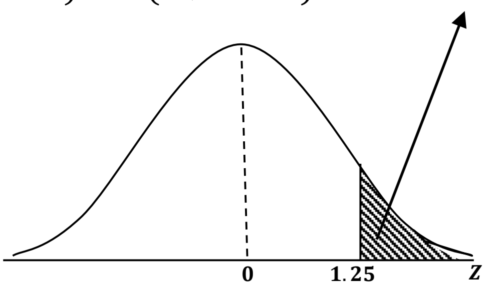
$$\mathbf{P(X > 71.25) = ?}$$

نحسب القيمة المعيارية المقابلة للقيمة **71.25** حيث:

$$\mathbf{z} = \frac{x - \mu_x}{\sigma_x} = \frac{\mathbf{71.25 - 65}}{\mathbf{5}} = \mathbf{1.25}$$

إذن:

$$\mathbf{P(}X\mathbf{ > 71.25) = P(Z > 1.25) = 0.50 - 0.3944 = 0.1056}$$



شكل (3−17)

(3–2–4) توزيع t :

يعتبر توزيع **t** من التوزيعات الاحتمالية المستمرة الهامة التي لها استخدامات كثيرة في موضوع الإحصاء الاستدلالي، ونستطيع تعريفه كما يلي:

تعريف توزيع t

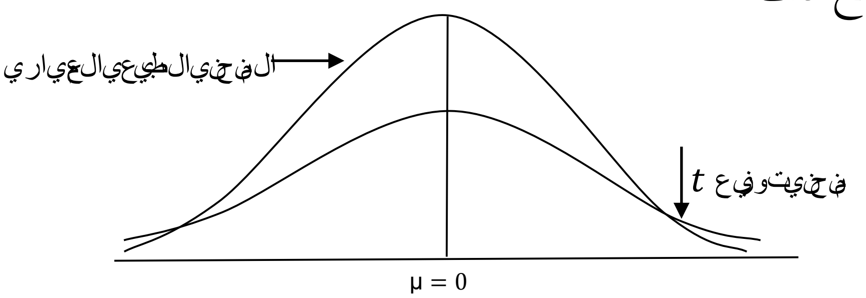
إذا كان z متغيراً عشوائياً يتبع التوزيع الطبيعي المعياري ، وأن Y متغير عشوائي يتبع توزيع χ^2 بدرجات حرية v ، فإن المتغير العشوائي التالي :

$$T = \frac{z}{\sqrt{Y/v}}$$

هو متغير عشوائي مستمر توزيعه الاحتمالي يطلق عليه توزيع **t** بدرجات حرية v.

تعتمد دالة كثافة الاحتمال لهذا التوزيع على معلمة واحدة وهي درجة الحريةv، ويرمز لتوزيع t عند درجة حرية v، بالرمز $t(v)$.

ومنحنى توزيع t يشبه منحنى التوزيع الطبيعي المعياري فهو ناقوسي الشكل ومتماثل حول وسطه الحسابي الذي يساوي الصفر، ولكن الفرق بينهما هو أن التوزيع الطبيعي المعياري تباينه يساوي الواحد الصحيح، بينما توزيع t تباينه أكبر من الواحد الصحيح، أي أن توزيع t أكثر تشتتا من التوزيع الطبيعي المعياري. وشكل (3–18) يوضح ذلك.



شكل (3–18)

ومن جدول (م.2) في ملحق الجداول الإحصائية نستطيع الحصول على قيمة t التي على يمينها مساحة قدرها α والتابعة لدرجة الحرية ν، فالعمود الأول في الجدول يمثل درجات الحرية ν والسطر الأول في الجدول يمثل المساحة التي على يمين القيمة t ، أما داخل الجدول فتوجد قيم t .

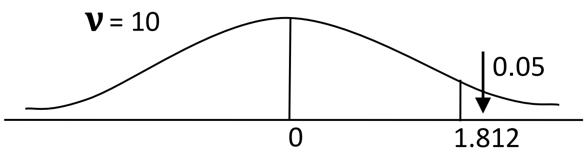
مثال (3-9) :

أوجد قيمة t التابعة لدرجة الحرية $\nu = 10$ والتي على يمينها مساحة قدرها **0.05**.

الحل:

نبحث في الجدول عن القيمة التي يتقاطع عندها العمود الذي يمثل مساحة قدرها **0.05** مع الصف الذي يمثل درجة حرية $\nu = 10$ فنجدها هي القيمة **1.812**، كما هو موضح في شكل (3−19)، ويعني ذلك أن احتمال أن يأخذ المتغير T عند درجة حرية **10** قيمة أكبر من **1.812** يساوي **0.05** أي أن:

$$P\left(t_{(10)} > 1.812\right) = 0.05$$



شكل (3−19)

مثال (3-10)

إذا كانت $t(\nu)$ ترمز لمتغير عشوائي يتبع توزيع t بدرجات حرية ν فباستخدام جدول (م.2) أوجد ما يلي :

1- $P(t_{(3)} > 4.541)$.

2- $P(t_{(11)} < 1.796)$.

3- $P(-1.753 \le t_{(15)} < 2.602)$.

الحل:

1- الاحتمال المطلوب $P(\ t\ _{(3)} > 4.541)$ هو عبارة عن المساحة على يمين القيمة **4.541**، عند درجة حرية **V = 3**، وبما أن جدول **(م.2)** يعطي المساحة على يمين القيمة، إذن بالبحث في هذا الجدول عن القيمة **4.541** في السطر الخاص بدرجة حرية **V = 3**، سنجد المساحة على يمين هذه القيمة مكتوبة أعلى القيمة في السطر الأول من الجدول وتساوي **0.01** ، أي أن :

$\boldsymbol{P\left(t_{(3)} > 4.541\right) = 0.01}$.

2-الاحتمال المطلوب$P(\ t_{(11)} < 1.796)$ هو عبارة عن المساحة على يسار القيمة **1.796**، وبما أن جدول **(م.2)** يعطى المساحة التي على يمين القيمة، فنوجد المساحة التي على يمين القيمة، ثم نطرحها من الواحد الصحيح (**لأن المساحة الكلية تحت دالة كثافة الاحتمال = 1**) فنحصل على المساحة التي على اليسار،

أي أن: $\boldsymbol{P\left(t_{(11)} < 1.796\ \right) = 1 - P(\ t_{(11)} \geq 1.796)}$

وبالبحث في جدول **(م.2)** عن القيمة **1.796** في السطر الخاص بدرجة حرية **v** تساوي **11** نجد المساحة على يمين هذه القيمة موجودة أعلى القيمة في السطر الأول من الجدول وتساوي **0.05** إذن الاحتمال المطلوب:

$$\boldsymbol{P\left(t_{(11)} < 1.796\ \right) = 1 - P\left(t_{(11)} \geq 1.796\right) = 1 - 0.05 = 0.95}$$

3-الاحتمال المطلوب$P(-1.753 \leq\ t_{(15)} < 2.602)$ يساوي المساحة المحصورة بين منحنى دالة كثافة الاحتمال لتوزيع t عند درجة حرية **=15** والمحور الأفقي وبين المستقيمين $\boldsymbol{t_{(15)}\ =\ 2.602}$، $\boldsymbol{t_{(15)}\ =\ -1.753}$ ، فمن جدول **(م.2)** نستطيع الحصول على المساحة على يمين القيمة **2.602**، حيث:

$\boldsymbol{P\left(t_{(15)} > 2.602\ \right) = 0.01}$

وكذلك نستطيع الحصول على المساحة على يمين القيمة **1.753**، حيث:

$$\boldsymbol{P\left(t_{(15)} > 1.753\right) = 0.05}$$

وبما أن منحنى توزيع t متماثل، فالمساحة على يمين القيمة **1.753** هي نفسها المساحة على يسار القيمة **1.753−**، أي أن:

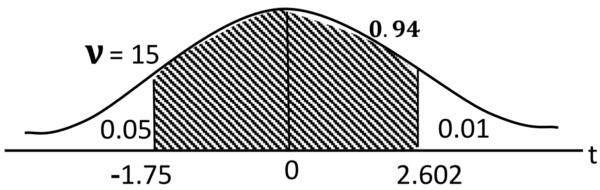
$$\boldsymbol{P\left(t_{(15)} < -1.753\right) = P\left(t_{(15)} > 1.753\right) = 0.05}$$

وبالتالي فالاحتمال المطلوب يساوي:

$$\boldsymbol{P\left(-1.753 \leq t_{(15)} < 2.602\right)}$$

$$\boldsymbol{= 1 - (0.01 + 0.05) = 0.94}$$

كما هو موضح في شكل **(3 − 20)**



شكل (3 − 20)

ملخص الفصل الثالث

درسنا في هذا الفصل توزيعين احتماليين متقطعين هامين، هما توزيع ذات الحدين وتوزيع بواسون، حيث دالة كتلة الاحتمال لتوزيع ذات الحدين هي:

$$f(\, x; n, p) = \, C_x^n P^x q^{n-x} \qquad x = 0, 1, 2, \dots \dots \dots n$$

P,n هما المعلمتان اللتان يعتمد عليهما توزيع ذات الحدين:

والوسط الحسابي والتباين لتوزيع ذات الحدين هما:

الوسط الحسابي $np = \mu$، **التباين** $npq = \sigma^2$

أما دالة كتلة الاحتمال لتوزيع بواسون فتأخذ الصيغة التالية:

$$f\,(x\,;\lambda\,) = \frac{e^{-\lambda}\ \lambda^x}{x!} \qquad x = 0, 1, 2, \dots \dots$$

يعتمد التوزيع على معلمة واحدة وهي λ

والوسط الحسابي للتوزيع = تباين التوزيع =λ

كذلك تطرقنا لأهم توزيع مستمر وهو التوزيع الطبيعي، ودالة كثافة الاحتمال لهذا التوزيع تأخذ الصيغة التالية:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}(\frac{x-\mu}{\sigma})^2} \qquad -\infty < x < \infty$$

ولحساب $P(x_1 < X < x_2)$ يجب تحويل $x_2\,.\,x_1$ إلى قيم معيارية $z_2\,.\,z_1$

حيث:

$$z_2 = \frac{x_2 - \mu_X}{\sigma_X}\,, z_1 = \frac{x_1 - \mu_X}{\sigma_X} \quad \text{و} \quad , \quad P(x_1 < X < x_2) = P(z_1 < Z < z_2)$$

كما درسنا توزيع مستمر آخر هام وله علاقة مباشرة بالتوزيع الطبيعي، وهو توزيع t.

تمارين (3-2)

1. إذا علمت أن المتغير العشوائي Z يتبع التوزيع الطبيعي المعياري، فأحسب ما يلي:

$P(Z < -1)$, $P(Z \leq 1.5)$, $P(Z > 1.5)$,
$P(Z \geq -2)$, $P(-1.0 \leq Z \leq 1.5)$

2. إذا علمت أن المتغير العشوائي Z يتبع التوزيع الطبيعي المعياري، فأحسب ما يلي، مع التوضيح بالرسم:

$P(Z < -1.23)$, $P(Z \leq 1.57)$, $P(Z > 1.23)$,
$P(Z \geq -2.75)$, $P(-1.45 \leq Z \leq 1.13)$, $P(Z = 1.32)$

3. إذا كان المتغير العشوائي X يتبع توزيعاً طبيعياً بوسط حسابي **9** وتباين **4**، فأحسب الاحتمالات التالية:

$P(X < 8)$, $P(X = 8)$, $P(8 \leq X \leq 11)$, $P(X > 10.5)$

4. إذا كان المتغير العشوائي X يتبع توزيعاً طبيعياً بوسط حسابي **110** وتباين **36**، فأحسب الاحتمالات التالية:

$P(X < 116)$, $P(95 \leq X \leq 120)$, $P(X > 115)$

5. إذا كانت أطوال طلبة المرحلة الثانوية تتوزع توزيعاً طبيعياً بمتوسط قدره **165** سم وتباين يساوي **9**. إذا اخترنا من طلبة هذه المرحلة طالباً واحداً عشوائياً فأحسب ما يلي:

أ. احتمال أن يكون طوله يتراوح بين **163** سم، **168** سم؟

ب. احتمال أن يكون طوله أقل من **162** سم؟

جـ. احتمال أن يكون طوله أكثر من **170** سم؟

6. إذا كان الدخل الشهري العائلي في مدينة ما، يتبع توزيعـاً طبيعيـاً بوسط حسابي قدره **250** دينار وانحراف معياري يساوي **20** دينار. إذا اخترنا من هذه المدينة عائلة واحدة عشوائيـاً، فأحسب ما يلي:

أ. احتمال أن يكون الدخل الشهري للعائلة أكثر من **300** دينـــــار؟

ب. احتمال أن يكون الدخل الشهري للعائلة أقل من **210** دينـــــار؟

ج. احتمال أن يتراوح الدخل الشهري للعائلة بين **260** و**220** دينـار؟

7. أوجد قيم $(\boldsymbol{t})$التالية:

أ. $\boldsymbol{t_{(12)}}$ التي على يسارها مساحة قدرها **0.99** .

ب. $\boldsymbol{t_{(11)}}$ التي على يمينها مساحة قدرها **0.05** .

ج. $\boldsymbol{t_{(20)}}$ التـي علـى يمينهـا مساحـــة **0.975** .

8. إذا كانت $\boldsymbol{t_{(v)}}$ ترمز لمتغير عشوائي يتبع توزيع$\boldsymbol{t}$ بدرجات حرية $\boldsymbol{v}$، فباستخدام جدول (م.**2**) أوجد ما يلي :

أ. $\boldsymbol{P(t_{(13)} > 1.771)}$.

ب. $\boldsymbol{P(t_{(21)} > 2.08)}$.

ج. $\boldsymbol{P(-1.833 \leq t_{(9)} < 2.821)}$.

الفصل الرابع
توزيعات المعاينة

(4-1) مقدمة:

البيانات الإحصائية هي المعلومات التي يجمعها الباحث عن الظاهرة التي يقوم بدراستها، وتشكل البيانات المادة الرئيسية في أي بحث إحصائي فعلي قدر صحتها تتوقف دقة البحث والتحليل الإحصائي، ويقوم الباحث بجمع البيانات باتباع أحد أسلوبين وهما:

(4-1-1) أسلوب الحصر الشامـل:

يتطلب أسلوب الحصر الشامل جمع البيانات عن كل أفراد المجتمع الإحصائي محل الدراسة. حيث المقصود بالمجتمع الإحصائي هو مجموعة كل المفردات التي يهتم بها موضوع البحث، وقد تكون هذه المفردات أشخاص أو أسر أو شركات أو حيوانات أو أشياء.

ومن أمثلة الحالات التي يستخدم فيها هذا الأسلوب هي التعدادات العامة للسكان حيث يتم جمع بيانات عن كل مفردة من مفردات المجتمع سواء كانت المفردة شخصاً أو أسرة أو مزرعة ... الخ.

وإذا كان مجتمعنا الإحصائي يشمل جميع طلاب جامعة ما مثلاً، فعند جمع البيانات باستخدام أسلوب الحصر الشامل، يجب جمع بيانات عن كل طالب من طلاب هذه الجامعة. وإذا كانت دراستنا خاصة بالدخل الشهري للعائلات القاطنة في مدينة ما، فالمجتمع الإحصائي يشمل كل العائلات القاطنة في هذه المدينة، وعند استخدام أسلوب الحصر الشامل يجب أن نجمع بيانات من كل عائلة من هذه العائلات.

(4-1-2) أسلوب المعاينة (أسلوب العينات):

المقصود بأسلوب المعاينة هو تجميع البيانات عن جزء فقط من مفردات المجتمع الإحصائي محل الاهتمام والدراسة، ويطلق على هذا الجزء مصطلح العينة. ونستطيع تعريف العينة كما يلي:

تعريف العينة

العينة هي جزء يتم اختياره من المجتمع محل الدراسة وذلك لغرض دراسة المجتمع من خلالها لأن دراسة المجتمع ككل غير ممكنة.

في اية دراسة إحصائية يجب أن يكون الهدف هو دراسة المجتمع ككل وليس دراسة العينة، ولكن تُستخدم العينة في الدراسة لأن الباحث لا يستطيع أن يجمع بيانات عن كل مفردات المجتمع محل الدراسة، وذلك للأسباب التالية:

(4-1-3) أسباب استخدام أسلوب المعاينة :

هناك أسباب تحتم علينا استخدام أسلوب العينات بدلاً من أسلوب الحصر الشامل، ويمكن تلخيص هذه الأسباب فيما يلي:

1- الإمكانيات المادية والفنية للباحث التي قد لا تسمح له بدراسة المجتمع بأكمله، والمقصود بالإمكانيات المادية والفنية هو المال المخصص للبحث والأشخاص المتدربين تدريباً جيداً على جمع البيانات.

2- عندما يكون المجتمع محل الدراسة مجتمعاً لا نهائياً، أي عدد مفرداته غير محدود فلا يستطيع الباحث جمع معلومات عن كل مفردة من مفرداته، مثل مجتمع الطيور والأسماك والحيوانات، ... الخ.

3-عندما يؤدي أسلوب الحصر الشامل إلى فناء المجتمع محل الدراسة وذلك عندما يؤدي فحص المفردات إلى هلاكها، فمثلاً عند فحص طلبية من البيض، فالبيضة هنا هي المفردة ولفحصها نضطر إلى كسرها مما يؤدي إلى إتلافها وبالتالي لا يمكن اتباع أسلوب الحصر الشامل، بل نكتفي بفحص جزء من هذه الطلبية أي نتبع أسلوب العينات.

4-توفير الوقت، فقد يستدعي الأمر أحياناً الحصول على نتائج البحث في وقت قصير بحيث يصعب اتباع أسلوب الحصر الشامل.

5-في حالة المجتمعات المتجانسة، أي تكون مفردات المجتمع متشابهة تماماً، فإن أسلوب الحصر الشامل يصبح إهداراً للوقت والجهد، فمثلاً يكفي اختيار قطعة من قماش الثوب بدلاً من الثوب كله إذا كان القماش متجانساً تماماً.

(4-2) عملية المعاينة:

هي عملية اختيار جزء من المجتمع محل الدراسة ويسمى هذا الجزء بالعينة للاستدلال على معالم وخواص المجتمع ككل، فهي عملية استنتاج إحصائي تقوم على التعميم من الجزء إلى الكل. والغاية الأساسية من إجراء عملية المعاينة هي تقدير القيم الحقيقية للمقاييس الإحصائية الخاصة بالمجتمع والتي يطلق عليها **"معالم"** من خلال بيانات العينة المختارة والمدروسة.

وتوجد عدة طرق لاختيار عينة من مجتمع بحيث تكون ممثلة له تمثيلاً سليماً، وباختلاف طريقة الاختيار تنتج أنواع مختلفة من العينات، وستعرض في هذا الكتاب إلى أهمها وهي العينة العشوائية البسيطة والتي يمكن تعريفها كما يلـــــــي:

العينة العشوائية البسيطة:

العينة العشوائية البسيطة هي العينة التي تسحب من المجتمع، بحيث يكون لكل مفردة من مفردات المجتمع نفس فرصة الظهور في العينة ، أي أن احتمال ظهور أية مفردة من مفردات المجتمع في العينة يكون متساوياً.

ولاختيار العينة بطريقة تضمن إعطاء نفس الفرصة لجميع مفردات المجتمع، يجب أن يكون الاختيار خاضعاً لعامل الصدفة المطلقة دون تدخل العامل البشري فيه، وهذا ما يطلق عليه الاختيار العشوائي، والذي تعرضنا لتعريفه في الفصل الأول. عندما تكون العينات عشوائية نستطيع استخدام الأساليب الإحصائية المختلفة ونظرية الاحتمالات لتحليل البيانات التي نحصل عليها من العينة، للاستدلال على معالم المجتمع الذي سحبت منه هذه العينة، وذلك بحساب مقاييس معينة من العينة يطلق عليها إحصاءة. وتستخدم الإحصاءة لتقدير معالم المجتمع المجهولة أو اتخاذ القرارات بخصوص صفات وخواص المجتمع الذي سحبت منه العينة. وتعرف الإحصاءة كما يلي:

تعريف الإحصاءة :

هي أي مقياس إحصائي تحسب قيمته من العينة المسحوبة من المجتمع محل الدراسة .

فمثلاً الوسط الحسابي للعينة عبارة عن إحصاءة ويرمز لها بالرمز $\overline{\mathbf{X}}$ وتباين العينة عبارة عن إحصاءة ويرمز له بالرمز $\mathbf{S}^2$ وهكذا ... وحيث أن قيمة الإحصاءة تعتمد على العينة المسحوبة، وبما أننا نستطيع أن نسحب أكثر من عينة من نفس المجتمع، فسنجد أن قيمة الإحصاءة ستتغير من عينة إلى أخرى، وبالتالي فإن

الإحصاءة عبارة عن متغير، وهذا هو الفرق الجوهري بين المعلمة والإحصاءة، **فالمعلمة عبارة عن قيمة ثابتة، بينما الإحصاءة عبارة عن متغير**. وبما أن الإحصاءة عبارة عن متغير فستكون متغيراً متقطعـًا أو متغيرا مستمرا، وسيكون لها توزيع احتمالي، والتوزيع الاحتمالي لأية إحصاءة يسمى توزيع معاينة.

تعريف توزيع المعاينة :

توزيع المعاينة هو التوزيع الاحتمالي لأية إحصاءة تحسب قيمها من كل العينات العشوائية ذات الحجم المتساوي والممكن سحبها من المجتمع.

فإذا سحبنا من المجتمع كل العينات العشوائية الممكنة ذات الحجم n ولكل عينة حسبنا قيمة الوسط الحسابي $\bar{X}$ فالتوزيع الاحتمالي للإحصاءة $\bar{X}$ يسمى توزيع المعاينة للوسط الحسابي $\bar{X}$ ، وإذا حسبنا لكل عينة قيمة التباين S^2 فالتوزيع الاحتمالي للإحصاءة S^2 يسمى توزيع المعاينة للتباين S^2 ، وهكذا ... فتوزيع المعاينة يوضح لنا نمط تغيير تلك الإحصاءات، وبالتالي نتمكن من إجراء استنتاج أو استدلال إحصائي حول القيم المناظرة لها في المجتمع.

سنعرض أمثلة توضح كيفية إيجاد توزيعات المعاينة وعلاقة هذه التوزيعات بتوزيع المجتمع الأصلي الذي سحبت منه العينات.

(3−4) توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$:

إذا كان لدينا مجتمع محدود حجمه N ووسطه الحسابي μ وتباينه σ^2، وسحبنا منه كل العينات العشوائية البسيطة الممكنة المتساوية في الحجم وليكن

حجمها n، وحسبنا الوسط الحسابي $\bar{X}$ لكل عينة ثم وضعنا هذه المتوسطات في جدول توزيع احتمالي، فهذا التوزيع الاحتمالي يطلق عليه توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$.

تعريف المعاينة للوسط الحسابي للعينة : هو التوزيع الاحتمالي للوسط الحسابي للعينة $\bar{X}$

مثال)4-1(:

شركة بها **5** أقسام، وفيما يلي عدد الموظفين في كل قسم:

16 , 12 , 10 , 8 , 4

أ) أوجد التوزيع الاحتمالي للمجتمع، ومنه أحسب الوسط الحسابي للمجتمع، وتباين المجتمع.

ب) فإذا اخترنا من هذا المجتمع كل العينات العشوائية التي تشمل قسمين (الاختيار مع عدم الإرجاع)، فأكتب توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$، وأحسب منه الوسط الحسابي والتباين للاحصاءة $\bar{X}$.

الحل :

أ) المجتمع الإحصائي (الشركة) يحتوي على **5** مفردات (أقسام)، والمتغير محل الدراسة هنا هو عدد الموظفين، فإذا رمزنا لهذا المتغير العشوائي المتقطع بالرمز X ، فبحساب دالة الاحتمال $f(x)$ لكل قيمة من قيم المتغير نحصل على توزيع المجتمع والموضح في جدول (4-1).

جدول (4-1): التوزيع الاحتمالي للمجتمع

x	4	8	11	12	16
$f(x)$	1/5	1/5	1/5	1/5	1/5

ومن جدول (4-1) نحسب الوسط الحسابي والتباين لهذا المجتمع كما يلي:

$$\mu = \sum x\, f(x) = (4)\left(\frac{1}{5}\right) + (8)\left(\frac{1}{5}\right) + (10)\left(\frac{1}{5}\right) + (12)\left(\frac{1}{5}\right) + (16)\left(\frac{1}{5}\right) = 10$$

$$\sigma^2 = \sum (x-\mu)^2\, f(x)$$

$$= (4-10)^2\left(\frac{1}{5}\right) + (8-10)^2\left(\frac{1}{5}\right) + (10-10)^2\left(\frac{1}{5}\right) + (12-10)^2\left(\frac{1}{5}\right) + (16-10)^2\,\frac{1}{5} = 16$$

ب- إذا سحبنا من هذا المجتمع مع عدم الإرجاع كل العينات الممكنة ذات الحجم $n = 2$ ، سيكون العدد الكلي للعينات الممكن سحبها هو:

$$\boxed{C_2^5 = \frac{5!}{2!\,3!} = 10}$$

فإذا حسبنا لكل عينة من هذه العينات العشرة، وسطها الحسابي $\overline{X}$، حيث $\overline{X}$ يساوي مجموع كل قيم العينة $\sum_{i=1}^{n} x_i$ مقسوما على حجم العينة n، أي أن:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

سنحصل على النتائج الموضحة في جدول **(4-2)**. ومن هذا الجدول نستطيع تكوين جدول توزيع المعاينة للوسط الحسابي للعينة $\overline{X}$ في حالة السحب مع عدم الإرجاع والموضح في جدول **(4-3)**.

جدول (4−2)

العينات ذات الحجم **2** الممكن سحبها من المجتمع مع عدم الإرجاع مقرونة بأوساطها الحسابية.

العينة	$\bar{X}$	العينة	$\bar{X}$
4,8	6	8,12	10
4,10	7	8,16	12
4,12	8	10,12	11
4,16	10	10,16	13
8,10	9	12,16	14

جدول (4-3): توزيع المعاينة للوسط الحسابي $\bar{X}$

$\bar{x}$	6	7	8	9	10	11	12	13	14
$f(\bar{x})$	1/10	1/10	1/10	1/10	2/10	1/10	1/10	1/10	1/10

ونستطيع حساب $\mu_{\bar{x}}$ و $\sigma^2_{\bar{x}}$ لتوزيع المعاينة من الصيغ التي استخدمناها عند حساب الوسط الحسابي وتباين المجتمع مع ملاحظة أن المتغير الذي نتعامل معه في توزيع المعاينة هو $\bar{X}$ بينما المتغير الذي نتعامل معه في حالة المجتمع هو X وبالتالي تكون الصيغ كما يلي:

$$\mu_{\bar{x}} = \sum \bar{x}\, f(\bar{x}) \quad , \quad \sigma^2_{\bar{x}} = \sum \left(\bar{x} - \mu_{\bar{x}}\right)^2 f\,(\bar{x})$$

والجدول التالي يوضح العمليات الحسابية اللازمة للحصول على هذين المقياسين.

$\bar{x}$	$f(\bar{x})$	$\bar{x}f(\bar{x})$	$(\bar{x} - \mu_{\bar{x}})^2$	$(\bar{x} - \mu_{\bar{x}})^2 f(\bar{x})$
6	**1/10**	**6/10**	**16**	**16/10**
7	**1/10**	**7/10**	**9**	**9/10**
8	**1/10**	**8/10**	**4**	**4/10**
9	**1/10**	**9/10**	**1**	**1/10**
10	**2/10**	**20/10**	**0**	**0**
11	**1/10**	**11/10**	**1**	**1/10**

12	1/10	12/10	4	4/10
13	1/10	13/10	9	9/10
14	1/10	14/10	16	16/10
		100/10		60/10

إذن: $\mu_{\bar{x}} = \sum \bar{x} f(\bar{x}) = \frac{100}{10} = 10$

$$\sigma^2_{\bar{x}} = \sum \left(\bar{x} - \mu_{\bar{x}}\right)^2 f\left(\bar{x}\right) = \frac{60}{10} = 6$$

نلاحـــــــظ أن: $\mu_{\bar{x}} = \mu$

بينمـــــــــا: $\sigma^2_{\bar{x}} \neq \sigma^2$

ولكن إذا حسبنا: $\frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

فسنجده يساوي تباين توزيع المعاينة $\sigma^2_{\bar{x}}$ ، حيث:

$$\frac{\sigma^2}{n} \times \frac{N-n}{N-1} = \frac{16}{2} \times \frac{5-2}{5-1} = 6$$

إذن: $\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$

ويطلق على المقدار $\frac{N-n}{N-1}$ معامل التصحيح.

وإذا كانت عملية السحبّ تمت مع الإرجاع، فسنجد أن العلاقة بين تباين توزيع المعاينة للوسط الحسابي للعينة وتباين المجتمع كما يلي:

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

وعندما يكون حجم المجتمع كبيراً جداً، أو عندما تكون نسبة حجم العينة إلى حجم المجتمع (n/N) أقل من أو تساوي 0.05، يؤول معامل التصحيح $\left(\frac{N-n}{N-1}\right)$ إلى الواحد الصحيح. وتكون العلاقة في حالة عدم الإرجاع هي نفسها في حالة الإرجاع وهي:

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n}$$

والنظرية التالية تلخص هذه العلاقات:

نظرية) 4- 1 (:

إذا كان لدينا مجتمع محدود، حجمه **N** ووسطه الحسابي μ وتباينه σ^2، وسحبنا منه كل العينات العشوائية الممكنة ذات الحجم **n** ، فإن الوسط الحسابي $\mu_{\bar{x}}$ والتباين $\sigma_{\bar{x}}^2$ لتوزيع المعاينة للوسط الحسابي للعينة يرتبط بالوسط الحسابي للمجتمع وتباين المجتمع حسب العلاقات التالية:

$$\mu_{\bar{X}} = \mu$$

في حالة السحب مع الإرجاع أو كان حجم المجتمع كبيراً أو $\frac{n}{N} \leq 0.05$:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n}$$

وفي حالة السحب مع عدم الإرجاع و $\frac{n}{N} > 0.05$:

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{N} \times \frac{N-n}{N-1}$$

(4-4): المعاينة من مجتمع يتبع التوزيع الطبيعي:

سندرس فيما يلي توزيع المعاينة للوسط الحسابي، عندما يكون المجتمع الذي سحبت منه العينات يتبع التوزيع الطبيعي، وذلك في حالتين:

(4-4-1) المجتمع يتوزع توزيعا طبيعيا بتباين σ^2 معلوم:

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا وسطه الحسابي μ وتباينه σ^2، وسحبنا منه كل العينات العشوائية الممكنة ذات الحجم **n**، فتوزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ سيتوزع توزيعـاً طبيعيـاً بوسط حسابي قدره $\mu_{\bar{x}}$ ، وتباين قدره $\sigma_{\bar{x}}^2$، إذن:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}}$$

ستتوزع توزيعا طبيعيًا معياريا، ويكون هذا صحيحًا بغض النظر عن حجم العينة كبيراً أم صغيراً.

ملاحظة:

يجب الانتباه هنا إلى أن المتغير الذي نحوله إلى المتغير المعياري **Z** هو المتغير $\bar{X}$ وبالتالي يجب أن نطرح منه وسطه الحسابي $\mu_{\bar{X}}$ ونقسم على انحرافه المعياري $\sigma_{\bar{X}}$

مثال)2-4(

إذا علمت أن أوزان مجتمع كبير جدا من الطلبة تتوزع توزيعًا طبيعيًا بوسط حسابي قدره **65** كيلو جرام وتباين قدره **25** فإذا سحبنا مع عدم الإرجاع من هذا المجتمع عينة بها **16** طالبًا، فما احتمال أن يكون الوسط الحسابي لأوزان هذه العينة أكبر من **67** كيلو جرام؟

الحل:

الاحتمال المطلوب: $P(\bar{X} > 67) = ?$

في هذه الدراسة المتغير محل الدراسة X هو الوزن، وبما أن مجتمع الأوزان تتوزع توزيعًا طبيعيًا، إذن توزيع المعاينة للوسط الحسابي للعينة سيكون توزيعًا طبيعيًا بغض النظر عن حجم العينة، وسيكون متوسط هذا التوزيع وتباينه كما يلي:

$$\sigma^2_{\bar{x}} = \frac{\sigma^2}{n} = \frac{25}{16} \quad , \quad \mu_{\bar{x}} = \mu = 65$$

عند حساب تباين توزيع المعاينة، لم نستخدم معامل التصحيح بالرغم من أن السحب تم مع عدم الإرجاع، لأن المجتمع كبير.

بما أن توزيع المعاينة توزيع طبيعي إذن لحساب الاحتمال المطلوب نحسب القيمة المعيارية المقابلة للقيمة $\bar{x} = 67$ وذلك كما يلي:

$$z = \frac{\bar{x} - \mu_{\bar{x}}}{\sigma_{\bar{x}}} = \frac{67 - 65}{\sqrt{\frac{25}{16}}} = \frac{2 \times 4}{5} = 1.60$$

وبالتالي فإن:

$$P(\bar{X} > 67) = P(Z > 1.60) = 0.50 - 0.4452 = 0.0548$$

مثال (3-4) :

إذا علمت أن درجات 100 طالباً في امتحان في مادة الإحصاء تتوزع توزيعاً طبيعياً بوسط حسابي قدره 70 وتباين قدره 9 ، فإذا سحبنا من هؤلاء الطلبة عينة عشوائية تشمل 25 طالباً حيث السحب تم مع عدم الإرجاع ، فما احتمال أن يكون الوسط الحسابي لهذه العينة أكبر من 71 ؟

الحل:

في هذا المثال يتكون المجتمع من كل الطلبة المشتركين في هذا الامتحان، أما المتغير العشوائي X محل الدراسة فهو درجة الطالب.

و المطلوب: $P(\bar{X} > 71) =?$

بما أن المجتمع يتوزع توزيعاً طبيعياً، فإن توزيع المعاينة للوسط الحسابي للعينة سيتوزع هو الآخر توزيعاً طبيعياً بمتوسط حسابي وتباين قدرهما على التوالي كما يلي:

$$\mu_{\bar{X}} = \mu = 70$$

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \times \frac{N-n}{N-1} = \frac{9}{25} \times \frac{100-25}{100-1} = 0.2727$$

هنا استخدمنا معامل التصحيح عند حساب التباين، وذلك لأن السحب تم مع عدم الإرجاع

$$\frac{n}{N}=\frac{25}{100}=0.25\ >0.05$$

ولحساب الاحتمال المطلوب نحسب القيمة المعيارية المقابلة للقيمة **71** فنجد أن:

$$z=\frac{\bar{X}-\mu_{\bar{X}}}{\sigma_{\bar{X}}}=\frac{71-70}{\sqrt{0.2727}}=1.92$$

$$P(\bar{X}>71)=P(Z>1.92)=0.50-0.4726=0.0274$$

(4-4-2) المجتمع يتوزع توزيعاً طبيعياً بتباين σ^2 مجهول:

علمنا مما سبق أنه إذا كانت العينات العشوائية مسحوبة من مجتمع يتوزع توزيعاً طبيعياً وسطه الحسابي μ وتباينه σ^2 ، فإن المتغير العشوائي **Z** حيث: $Z=\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}$ سيتوزع توزيعاً طبيعياً معيارياً.

أما إذا كان تباين المجتمع σ^2 مجهولاً، فنستعمل تباين العينة S^2 كتقدير لتباين المجتمع المجهول σ^2 ، حيث تباين العينة يحسب كما يلي :

$$S^2=\frac{\sum(X-\bar{X})^2}{n-1}$$

وسنحصل على المتغير العشوائي التالي:

$$\boxed{T=\frac{\bar{X}-\mu}{\sqrt{S^2/n}}}$$

وهذا المتغير يتبع توزيع **t** بدرجات حرية **(n-1)**، وقد قمنا بدراسة هذا التوزيع في الفصل الثالث .

مثال (4-4) :

إذا كان المتغير العشوائي **X** يمثل أوزان علب نوع معين من العصير التي ينتجها أحد المصانع، وأن هذا المتغير يتبع التوزيع الطبيعي بوسط حسابي يساوي **100**

جرام. سُحبت عينة من إنتاج هذا المصنع حجمها 9 علب، وكان تباينها يساوي 16، ما هو احتمال أن يزيد الوسط الحسابي لهذه العينة عن 103.85 جرام؟

الحل:

نلاحظ هنا ان تباين المجتمع σ^2 مجهول، وحجم العينة صغير (أقل من 30)

وأن المتغير العشوائي X يتبع التوزيع الطبيعي إذن :

$$T=\frac{\bar{X}-\mu}{\sqrt{S^2/n}}$$

يتبع توزيع t بدرجات حرية (n-1)، حيث:

$$(n-1)=9-1=8$$

$$T=\frac{\bar{X}-\mu}{\sqrt{S^2/n}}=\frac{103.85-100}{\sqrt{16/9}}=2.89$$

إذن الاحتمال المطلوب:

$$P(\bar{X}>103.85)=P\left(t_{(8)}>2.89\right)=0.01$$

مثال (5-4) :

إذا كانت قيمة أرصدة الحسابات الجارية في أحد المصارف تتبع توزيعاً طبيعياً بوسط حسابي يساوي 20 ألف دينار، فإذا اخترنا عينة عشوائية تشمل 25 حساباً، وعلمت أن تباينها يساوي 36، فما احتمال أن يكون الوسط الحسابي للأرصدة الجارية التي تضمها العينة:

أ) أقل من 17.50 ألف دينار.

ب) أقل من 21.50 ألف دينار.

الحل:

نلاحظ هنا أن تباين المجتمع σ^2 مجهول، وحجم العينة صغير (أقل من 30)

وأن المتغير العشوائي X يتبع التوزيع الطبيعي إذن:

$$T=\frac{\bar{X}-\mu}{\sqrt{S^2/n}}$$

يتبع توزيع **t** بدرجات حرية **(n-1)**، حيث:

$$(n-1) = 25 - 1 = 24$$

أ- لحساب احتمال أن يكون الوسط الحسابي للعينة أقل من **17.50** ألف دينار، أي حساب $P(\bar{X} < 17.50)$ ، نحسب قيمة **t** المقابلة لقيمة $\bar{X} = 17.50$ ، وذلك كما يلي :

$$T = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}} = \frac{17.50 - 20.00}{\sqrt{\frac{36}{25}}} = -2.08$$

وبما أن منحنى توزيع **t** متماثل، إذن:

$$P(\bar{X} < 17.50) = P\left(t_{(24)} < -2.08\right) = P\left(t_{(24)} > 2.08\right) \cong 0.025$$

ب- لحساب احتمال أن يكون الوسط الحسابي للعينة أقل من **21.50** ألف دينار، أي حساب $P(\bar{X} < 21.50)$ ، نحسب قيمة **t** المقابلة لقيمة $\bar{X} = 21.50$ ، وذلك كما يلي:

$$T = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}} = \frac{21.50 - 20.00}{\sqrt{\frac{36}{25}}} = 1.25$$

$$\therefore P(\bar{X} < 21.50) = P(t_{(24)} < 1.25)$$

بالبحث في جدول (م.2)، مقابل درجة الحرية **24**، عن أقرب قيمة للقيمة **1.25**، فسنجدها هي القيمة **1.318**، وبما أن الجدول يتعامل مع المساحة التي على يمين القيمة، والمساحة الكلية تحت المنحنى تساوي **1** ، إذن:

$$P\left(t_{(24)} < 1.318\right) = 1 - P\left(t_{(24)} > 1.318\right) = 1 - 0.10 = 0.90$$

إذن الاحتمال المطلوب:

$$P(\bar{X} < 21.50) = P\left(t_{(24)} < 1.25\right) \cong 0.90$$

(4–5) المعاينة من مجتمع لا يتبع التوزيع الطبيعي:

إذا كان حجم العينة كبيراً (أكبر من 30)، فإن توزيع المعاينة للوسط الحسابي للعينة نحصل عليه باستخدام نظرية النهاية المركزية. حيث منطوق هذه النظرية كما يلي:

نظرية (4–2) (نظرية النهاية المركزية):

إذا كان لدينا مجتمع وسطه الحسابي μ وتباينه σ^2 ، وسحبنا منه كل العينات العشوائية الممكنة ذات الحجم الكبير (n أكبر من أو يساوي 30) فتوزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ سيكون توزيعاً قريباً من التوزيع الطبيعي بمتوسط قدره $\mu_{\bar{X}}$ وتباين قدره $\sigma^2_{\bar{X}}$ وعليه فإن

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$$

ستتوزع توزيعاً قريباً من التوزيع الطبيعي المعياري .

ونظرية النهاية المركزية لها أهمية كبيرة في الاستدلال الإحصائي، فباستخدامها نعتبر توزيع المعاينة للوسط الحسابي للعينة توزيعا قريبـًا من التوزيع الطبيعي، وبالتالي نستطيع تطبيق خواص التوزيع الطبيعي دون الحاجة إلى معرفة التوزيع الاحتمالي للمجتمع طالما أن حجم العينة كبير $(n \geq 30)$.

مثال (4-6) :

إذا علمت أن الوسط الحسابي لدرجات امتحان في الإحصاء، أشترك فيه عدد كبير من الطلبة، يساوي 73 وبتباين 25، فإذا اخترنا من الطلبة الذين اشتركوا في هذا الامتحان عينة عشوائية تشمل 100 طالب (حيث السحب تم مع عدم الإرجاع)، فما احتمال أن يكون الوسط الحسابي لدرجات الطلبة الذين تشملهم هذه العينة أكثر من 74؟

الحل:

بما أن، توزيع المجتمع ليس توزيعا طبيعياً، ولكن حجم العينة كبير $(n > 30)$ ، فإن توزيع المعاينة للوسط الحسابي للعينة $\overline{X}$ هو توزيع قريب من التوزيع الطبيعي (نظرية النهاية المركزية) بمتوسط وتباين قدرهما على التوالي:

$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} = \frac{25}{100} \qquad , \qquad \mu_{\overline{x}} = \mu = 73$$

عند حساب تباين توزيع المعاينة لم نستخدم معامل التصحيح بالرغم أن السحب تم مع عدم الإرجاع، لأن حجم المجتمع كبير.

ولحساب الاحتمال المطلوب نحسب القيمة المعيارية المقابلة للقيمة **(74)** وذلك كما يلي :

$$Z = \frac{\overline{x} - \mu_{\overline{x}}}{\sigma_{\overline{x}}} = \frac{74 - 73}{\sqrt{\frac{25}{100}}} = \frac{1}{0.50} = 2.00$$

وبالتالي فإن الاحتمال المطلوب:

$$P(\overline{X} > 74) = P(Z > 2.00) = 0.50 - 0.4772 = 0.0228$$

ملخص الفصل الرابع

علمنا أن الاحصاءة هي أي مقياس إحصائي نحسب قيمته من العينة المسحوبة من المجتمع محل الدراسة، وبما أن قيمة الاحصاءة تتغير من عينة عشوائية إلى أخرى، فهي متغير عشوائي، ولها توزيع احتمالي يطلق عليه توزيع المعاينة، وقد درسنا توزيع المعاينة للوسط الحسابي للعينة بإسهاب ووجدنا أن هناك علاقة تربط الوسط الحسابي وتباين توزيع المعاينة للوسط الحسابي للعينة مع الوسط الحسابي وتباين توزيع المعاينة للوسط الحسابي للعينة مع الوسط الحسابي وتباين المجتمع الذي سحبت منه العينات، حيث:

$$\mu_{\bar{x}} = \mu$$

في حالة السحب مع الإرجاع أو كان حجم المجتمع كبيرا: $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$

في حالة السحب مع عدم الإرجاع وكان حجم المجتمع صغيرا:

$$\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n} \times \frac{N-n}{N-1}$$

وعرفنا أنه إذا كان توزيع المجتمع الذي سحبت منه العينات يتوزع توزيعاً طبيعياً بتباين معلوم، فإن توزيع المعاينة للوسط الحسابي للعينة سيتبع هو الآخر توزيعاً طبيعياً، أما إذا كان توزيع المجتمع ليس توزيعاً طبيعياً، فإن توزيع المعاينة للوسط الحسابي للعينة سيتوزع توزيعاً قريباً من التوزيع الطبيعي إذا كان حجم العينة كبيراً، أي $n \geq 30$ **(نظرية النهاية المركزية)**. ويكون المتغير العشوائي التالي:

$$Z = \frac{\bar{X} - \mu_{\bar{X}}}{\sigma_{\bar{X}}}$$

يتوزع توزيعا قريباً من التوزيع الطبيعي المعياري.

وعندما يتوزع المجتمع توزيعاً طبيعياً بتباين مجهول، فنستعمل تباين العينة s^2 كتقدير لتباين المجتمع المجهول σ^2 ، ونحصل على المتغير العشوائي التالي:

$$T = \frac{\bar{X} - \mu}{\sqrt{s^2/n}}$$

وهـذا المتغيـر يتبـع توزيـع $\mathbf{t}$ بدرجـات حريـــة $(n-1)$.

تمارين (4)

1- تكلم عن أساليب جمع البيانات الإحصائية.

2- ما أسباب استخدام أسلوب المعاينـــــة؟

3- أذكر الفرق بين المعلمة والإحصـــــاءة .

4- أذكر منطوق نظرية النهاية المركزية، ومدى أهمية هذه النظرية.

5- يتكون مجتمع من القيم التالية: **14,6,2**

أ) أوجد التوزيع الاحتمالي للمجتمع، ومنه أحسب الوسط الحسابي للمجتمع، وتباين المجتمع.

ب) إذا اخترنا من هذا المجتمع كل العينات العشوائية التي تشمل مفردتين (الاختيار مع عدم الإرجاع)، فأوجد توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ وأحسب منه الوسط الحسابي والتباين للإحصاءة $\bar{X}$.

جـ) تحقق من صحة العلاقة التي تربط بين الوسط الحسابي للمجتمع μ والوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة، $\mu_{\bar{X}}$ ومن صحة العلاقة التي تربط بين تباين المجتمع σ^2 وتباين توزيع المعاينة للوسط الحسابي للعينة $\sigma_{\bar{X}}^2$.

6- يتكون مجتمع من القيم التالية: **18,12,9,3**

أ) أوجد التوزيع الاحتمالي للمجتمع، ومنه أحسب الوسط الحسابي للمجتمع، وتباين المجتمع.

ب) إذا اخترنا من هذا المجتمع كل العينات العشوائية التي حجمها 3، (الاختيار مع عدم الإرجاع)، فأوجد توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$. وأحسب منه الوسط الحسابي والتباين للإحصاءة $\bar{X}$.

ج) تحقق من صحة العلاقة التي تربط بين الوسط الحسابي للمجتمع μ والوسط الحسابي لتوزيع المعاينة للوسط الحسابي للعينة $\mu_{\bar{X}}$، ومن صحة العلاقة التي تربط بين تباين المجتمع σ^2 وتباين توزيع المعاينة للوسط الحسابي للعينة $\sigma_{\bar{X}}^2$.

-7 إذا كان التوزيع الاحتمالي للمتغير العشوائي كما يلي:

x	4	5	7	11	12
$f(x)$	3.1	3.3	3.2	3.1	3.3

إذا سحبنا من هذا المجتمع كل العينات العشوائية ذات الحجم n = 4 مع الإرجاع، فأوجد الوسط الحسابي والتباين والانحراف المعياري لتوزيع المعاينة للوسط الحسابي للعينة $\bar{X}$.

-8 إذا كان لدينا مجتمع وسحبنا منه مع الإرجاع كل العينات العشوائية ذات الحجم n = 5 فوجدنا أن توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ كما يلي:

$\bar{x}$	3	4	7	13	11
$f(\bar{x})$	3.35	3.13	3.15	3.33	0.40

أحسب الوسط الحسابي والتباين والانحراف المعياري لهذا المجتمع.

-9 إذا سحبنا عينة عشوائية حجمها (16) مع الإرجاع من مجتمع يتوزع توزيعاً طبيعياً وسطه الحسابي 100 وتباينه 49.

أ) أحسب احتمال أن يقع الوسط الحسابي للعينة بين 97 , 103

10. إذا كانت أوزان **1500** طالب بمدرسة ثانوية تتبع توزيعاً طبيعياً بوسط حسابي قدره **35** كيلو جرام وتباين قدره **16**، إذا اخترنا من هذه المدرسة مع عدم الإرجاع، عينة عشوائية تشمل **100** طالب، فأحسب ما يلي:

أ) احتمال أن يكون الوسط الحسابي لهذه العينة أكثر **من 34** كيلو جـــــرام.

ب) احتمال أن تتراوح قيمة الوسط الحسابي للعينة بين **34** ، **36** كيلو جرام؟

11. إذا كانت أطوال **1000** طالب تتوزع توزيعا طبيعيا بوسط حسابي قدره **163.5** سم وتباين قدره **25**، إذا اخترنا منهم مع عدم الإرجاع، عينة عشوائية تشمل **64** طالبا، فأحسب ما يلي:

أ) احتمال أن يكون الوسط الحسابي لهذه العينة أقل من **162** سم.

ب) احتمال أن يتراوح الوسط الحسابي للعينة بين **162 و 165** ســم.

12. مصنع للمواد الغذائية أنتج **10000** قطعة حلوى من نوع معين، وكان متوسط أوزان هذه القطع المنتجة $\mu = 41$ جراما، فإذا سحبنا من إنتاج هذا المصنع عينة عشوائية تحتوى على **9** قطع وكان تباينها يساوى **1.44** جراما فأحسب احتمال أن يكون الوسط الحسابي للعينة أكثر من **42.3** جراما.

13. مصنع به **500** عامل، وكان عدد الوحدات المنتجة من قبل العامل يوميا يتبع توزيعا طبيعيا بوسط حسابي قدره **15** وحدة، فإذا اخترنا عينة عشوائية تشمل **16** عاملاً من هذا المصنع مع عدم الإرجاع، وكان تباين هذه العينة يساوى **4**، فما احتمال أن يكون وسطها الحسابي أقل من **14**؟

14. إذا كان المتغير العشوائي **X** وسطه الحسابي **110** وتباينه **100**، سحبت منه مع الإرجاع عينه عشوائية حجمها **64**، أحسب احتمال أن يزيد الوسط الحسابي لهذه العينة عن **113**.

15. قرية بها **800** عائلة، وكان الوسط الحسابي للدخل الشهري لهذه العائلات **210** دينار شهريا، بانحراف معياري **50** دينار، فإذا اخترنا من هذه القرية **100** عائلة (**مع عدم الإرجاع**) فما احتمال أن يكون الوسط الحسابي للدخل الشهري لهذه العينة أكثر من **215** دينار.

الفصل الخامس
التقدير الإحصائي

(5–1) مقدمـــــــة:

في أغلب الدراسات لا نستطيع معرفة القيمة الحقيقية لمعلمة المجتمع محل البحث، وذلك لأنه لكي نحسب قيمة معلمة يجب أن يكون لدينا بيانات عن كل مفردات المجتمع دون استثناء، ولكن في أغلب الدراسات لا يمكننا جمع بيانات عن كل مفردات المجتمع. فمن أهم المواضيع التي يهتم بها الاستنتاج الإحصائي هي كيفية تقدير معالم المجتمع المجهولة (المعالم المجهولة للتوزيعات الاحتمالية)، مثل الوسط الحسابي للمجتمع أو تباين المجتمع أو أي مقياس إحصائي آخر خاص بالمجتمع، باستخدام بيانات نتحصل عليها عينة عشوائية مسحوبة من ذلك المجتمع.

(5–2) أنواع التقدير:

يوجد نوعان من التقدير الإحصائي للمعالم المجهولة هما:

1. **التقدير بقيمة.**
2. **التقدير بفتـرة.**

وسنقوم بعرض كل منهما فيما يلي.

(5–2–1) التقدير بقيمة:

المقصود بهذا النوع من التقدير هو تقدير معلمة المجتمع المجهولة بإحصاءة نحسب قيمتها من بيانات عينة عشوائية مسحوبة من هذا المجتمع، أي نقوم بتقدير المعلمة بقيمة واحدة فقط تحسب من العينة ولذلك يسمى التقدير بقيمة، فمثلا نقدر الوسط الحسابي للمجتمع μ بالوسط الحسابي للعينة $\bar{X}$، ونقدر تباين المجتمع σ^2 بتباين العينة S^2 وهكذا ...

مثال (5-1) :

إذا سحبنا عينة عشوائية تشمل 10 عائلات من العائلات القاطنة في مدينة ما، وكان الإنفاق الشهري بالدينار لكل عائلة من هذه العائلات كما يلي:

120 ، 165 ، 263 ، 252 ، 175 ، 168 ، 232 ، 255 ، 200 ، 240

باستخدام بيانات هذه العينة قدِّر الوسط الحسابي للإنفاق الشهري لكل العائلات القاطنة في هذه مدينة.

الحل:

في هذه الدراسة نجد أن، المجتمع هو كل العائلات القاطنة في هذه مدينة، والمطلوب تقدير الوسط الحسابي لإنفاق هذه العائلات، أي المطلوب تقدير الوسط الحسابي للمجتمع μ. فنستطيع تقدير المعلمة المجهولة وهي الوسط الحسابي للمجتمع **بقيمة واحدة** تحسب من العينة وهي الوسط الحسابي للعينة $\overline{X}$ حيث:

$$\overline{X} = \frac{\sum x}{n} = \frac{2070}{10} = 207$$

إذن نقدر الوسط الحسابي للإنفاق الشهري للعائلات القاطنة في هذه مدينة بالقيمة **207** دينار، وهذه القيمة ليست هي القيمة الحقيقية للوسط الحسابي للمجتمع وإنما هي تقدير لهذا الوسط، ونرمز للقيمة المقدّرة للوسط الحسابي للمجتمع بالرمز $\hat{\mu}$ أي أن :

$$\hat{\mu} = 207$$

الإشارة ^ تعنى في علم الإحصاء القيمة المقدَّرة.

وأحيانا نجد أكثر من إحصاءة يمكن استخدام قيمتها كتقدير للمعلمة المجهولة، لذلك توجد معايير تساعدنا على اختيار الإحصاءة التي تعتبر أفضل من غيرها لتقدير المعلمة المجهولة، أي توجد خصائص يجب أن تتوفر في المقدَّر حتى يكون مقدِّرا جيدا، ولكن لن نتعرض لها في هذا الكتاب.

(5-2-2) التقدير بفترة:

بدلا من تحديد قيمة واحدة تستخدم لتقدير المعلمة المجهولة، فإننا في هذا النوع من التقدير نحدد فترة معينة تقع فيها المعلمة المجهولة، فمثلا إذا رمزنا للمعلمة المجهولة بالرمز θ حيث θ قد تكون الوسط الحسابي μ أو التباين σ^2 أو النسبة P أو أي مقياس إحصائي آخر خاص بالمجتمع ، فإننا نحدد فترة تقع فيها هذه المعلمة كما يلي :

$$\boxed{\hat{\theta} - E \leq \theta \leq \hat{\theta} + E}$$

حيث:

$\hat{\theta}$ **:** الإحصاءة المستخدمة كأفضل مقدِّر بالقيمة.

E : قيمة نعتمد في حسابها على بيانات العينة وحجم العينة ومستوى الثقة في التقدير.

وتسمى هذه الفترة بفترة الثقة للمعلمة θ ويسمى المقدار $(\hat{\theta} - E)$ بالحد الأدنى لفترة الثقة والمقدار $(\hat{\theta} + E)$ بالحد الأعلى لفترة الثقة .

وإذا كان معامل الثقة $(1 - \alpha)$ حيث $0 < \alpha < 1$ فيعنى ذلك أن :

$$\boxed{p(\hat{\theta} - E \leq \theta \leq \hat{\theta} + E) = 1 - \alpha}$$

والمقصود بذلك أن احتمال احتواء الفترة $(\hat{\theta} - E$ و $\hat{\theta} + E)$ على المعلمة المجهولة θ هو **$(1-\alpha)$.**

فمثلا إذا كان **$1 - \alpha = 0.95$** فيعنى ذلك أنه إذا سحبنا كل العينات العشوائية من المجتمع (أو عددا كبيرا جدا من العينات العشوائية من المجتمع) وحسبنا لكل عينة فترة الثقة، فسنجد أن **95%** من هذه الفترات تحتوى على المعلمة المجهولة θ وأن **5%** من الفترات لا تحتوى على θ وتسمى النسبة **95% بمستوى الثقة** .

وسنقوم فيما يلي باستعراض كيفية حساب فترة الثقة للوسط الحسابي للمجتمع.

(3-5) فترة الثقة للوسط الحسابي للمجتمع μ :

(1-3-5) عندما يكون تباين المجتمع σ^2 معلوما:

علمنا من دراستنا السابقة وفقا لنظرية النهاية المركزية، أنه إذا كان لدينا مجتمع وسطه μ وتباينه σ^2 (ليس من الضروري أن يكون توزيعه توزيعا طبيعيا) ، وسحبنا منه كل العينات العشوائية الممكنة ذات الحجم n بحيث n تكون كبيرة بما فيه الكفاية $(n > 30)$ فإن توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ سيكون قريبا من التوزيع الطبيعي بوسط حسابي وتباين قدرهما :

$$\sigma_{\bar{x}}^2 = \frac{\sigma^2}{n} \quad ، \mu_{\bar{X}} = \mu$$

أما إذا كان المجتمع يتوزع توزيعا طبيعيا، فإن توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ سيكون توزيعا طبيعيا، وذلك سواء كان حجم العينة n صغيرا أو كبيرا، وفي هذه الحالات المذكورة نستطيع تحويل $\bar{X}$ إلى المتغير المعياري Z حيث:

$$\boxed{Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2 | n}}}$$

ويكون توزيع المتغير Z طبيعيا معياريا.

فمثلا نستطيع من جدول (م.**1**) الحصول على قيمة Z التي على يمينها مساحة قدرها **0.025**، وبالتالي المساحة بينها وبين الصفر تساوى **0.4750** فسنجدها مقابلة للقيمة **1.96**، كذلك نستطيع الحصول على القيمة التي على يسارها في المساحة **0.025**، فسنجدها مقابلة للقيمة **1.96-**، ونلاحظ أن القيمة المطلقة للقيمتين متساوية والاختلاف بين القيمتين في الإشارة فقط، وبصفة عامة في حالة التوزيع الطبيعي المعياري إذا كانت لدينا قيمتان متساويتان في القيمة المطلقة ومختلفتان في الإشارة

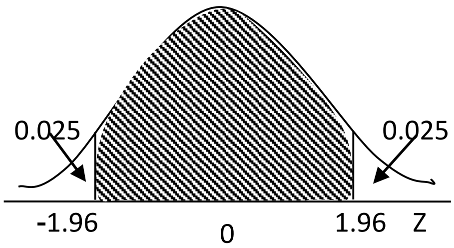
فقط، فستكون المساحة التي على يمين القيمة الموجبة تساوى المساحة التي على يسار القيمة السالبة ، لأن المنحنى الطبيعي المعياري متماثل.

وبما أن المساحة على يمين القيمة **1.96** تساوي **0.025** والمساحة على يسار القيمة **1.96−** تساوي **0.025** ، إذن المساحة بين القيمتين ستكون :

$$\mathbf{1-(0.025+0.025)=0.95}$$

وشكل **(5−1)** يوضح ذلك، وحيث أن المساحة تحت المنحنى الاحتمالي هي عبارة عن احتمالات، فيعنى ذلك أن احتمال أن يأخذ المتغير Z قيمة محصورة بين القيمتين **1.96−** و **1.96** يساوى **0.95** ونعبر عن ذلك كما يلي :

$$\boldsymbol{P(-1.96 \le Z \le 1.96)=0.95}$$



شكل (5−1)

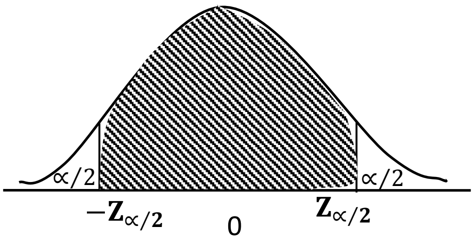
إذن نستطيع التعبير عن الاحتمال السابق كما يلي:

$$\boldsymbol{P\left(-1.96 \le \frac{\overline{X}-\mu}{\sqrt{\sigma^2|n}} \le 1.96\right)=0.95}$$

وبصفة عامة إذا رمزنا للمساحة **0.95** بالرمز **α −1** وبالتالي ستكون المساحة **0.025** مساوية للمساحة $\alpha/2$ وتكون القيمة **1.96** هي قيمة **Z** التي على يمينها مساحة $\alpha/2$ ونرمز لهذه القيمة بالرمز $Z\alpha_{/2}$، بالطبع سنرمز للقيمة **1.96−** بالرمز ، $-Z\alpha_{/2}$ وهكذا نستطيع التعبير عن الاحتمال السابق بصفة عامة كما يلي :

$$\boxed{\boldsymbol{P\left(-Z\alpha_{/2} \le \frac{\overline{X}-\mu}{\sqrt{\sigma^2|n}} \le Z\alpha_{/2}\right)=1-\alpha}}$$

وشكل (5−2) يوضح ذلك.



شكل (5−2)

وبإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة الموجودة بين القوسين بحيث نجعل الحد الأوسط للمتباينة يحتوى على المعلمة المجهولة $\boldsymbol{\mu}$ فقط ، نحصل على :

$$P\left(\ \bar{X} - Z_{\frac{\alpha}{2}}\sqrt{\sigma^2|n}\ \leq \mu \leq \bar{X} + Z_{\alpha/2}\sqrt{\sigma^2/n}\ \right) = 1 - \alpha$$

ويعنى ذلك أن احتمال وقوع الوسط الحسابي للمجتمع μ (المجهول) بين القيمة $(\bar{X} - Z_{\alpha/2}\sqrt{\sigma^2|n})$ والقيمة $(\bar{X} + Z_{\alpha/2}\sqrt{\sigma^2|n})$ يساوى $(1-\alpha)$.

- ويسمى المقدار $\bar{X} - Z_{\alpha/2}\sqrt{\sigma^2|n}$ بالحد الأدنى لفترة الثقة .
- ويسمى المقدار $\bar{X} + Z_{\alpha/2}\sqrt{\sigma^2|n}$ بالحد الأعلى لفترة الثقــة .

أما الاحتمال $(1-\alpha)$ فيسمى معامل الثقة وعادة نعبر عنه بنسبة مئوية، أي يكتب $100\%\ (1-\alpha)$ ويسمى مستوى الثقة، وقد تكون أية نسبة ولكن النسب الدارجة الاستعمال **90%، 95%، 99%** .

وبالتالي فإن فترة الثقة للوسط الحسابي للمجتمع μ عندما يكون تباين المجتمع معلوما وعند مستوى ثقة $100\%(1-\alpha)$ **هي :**

$\bar{X} - Z_{\alpha/2}\sqrt{\sigma^2|n}$ ، $\bar{X} + Z_{\alpha/2}\sqrt{\sigma^2|n}$

مثال (2-5):

إذا علمت أن تباين أوزان كل الطلبة المسجلين في إحدى الجامعات في سنة معينة يساوى **144**، وسحبنا من هؤلاء الطلبة عينة عشوائية تشمل **100** طالب، ووجدنا أن الوسط الحسابي لأوزانهم يساوى **64** كيلو جرام، فمن هذه البيانات قدر الوسط الحسابي لأوزان كل طلبة المسجلين في هذه الجامعة تلك السنة، وذلك باستخدام فترة ثقة بمستوى قدرة **99%**.

الحـل:

المجتمع في هذه الدراسة يتكون من أوزان كل الطلبة المسجلين في هذه الجامعة في تلك السنة، وتباينه معلوم، وهنا بالرغم من عدم ذكر توزيع المجتمع فنستطيع استعمال المتغير الطبيعي المعياري Z للحصول على فترة الثقة المطلوبة، لأن حجم العينة كبير ($n>30$) وذلك وفقا لنظرية النهاية المركزية. والفترة هي:

$$\left(\bar{X} - Z_{\alpha/2}\sqrt{\sigma^2|n} \quad , \quad \bar{X} + Z_{\alpha/2}\sqrt{\sigma^2|n}\right)$$

حيث: $(1-\alpha)100\% = 99\%\ \bar{X} = 64\ \ n = 100\ ,\ \ \sigma^2 = 144$

إذن: $1-\alpha = 0.99$ ، $\alpha = 0.01$ ، $\alpha/2 = 0.005$ ، $Z_{\alpha/2} = Z_{0.005} = 2.58$

وبالتعويض في الحد الأدنى والحد الأعلى لفترة الثقة نحصل على:

الحد الأدنى لفترة الثقة:

$$\bar{X} - Z_{\alpha/2}\sqrt{\sigma^2/n} = 64 - (2.58)\sqrt{144/100} = 64 - 3.096 = 60.904$$

الحد الأعلى لفترة الثقة:

$$\bar{X} + Z_{\alpha/2}\sqrt{\sigma^2/n} = 64 + (2.58)\sqrt{144/100} = 64 + 3.096 = 67.096$$

إذن فترة الثقة للوسط الحسابي لأوزان كل الطلبة المسجلين في هذه الجامعة في تلك السنة، عند مستوى ثقة **99%** هي:

(60.904 ، 67.096)

أي نستطيع القول بثقة **99%** بأن الوسط الحسابي الحقيقي لأوزان كل الطلبة المسجلين في تلك السنة يقع بين **60.904** كيلو جرام **و 67.096** كيلو جرام .

مثال (3-5) :

إذا كان المبلغ المودع في الحسابات الجارية في أحد المصارف، يتبع توزيعا ما بانحراف معياري قدره **3150** دينار، فإذا اخترنا من هذا المصرف عينة عشوائية تشمل **100** حساب جارى، ووجدنا أن الوسط الحسابي للمبالغ المودعة في الحسابات التي تشملها العينة **8525** دينار. فقدر الوسط الحسابي للمبالغ المودعة في كل الحسابات الجارية في هذا المصرف، وذلك باستخدام فترة ثقة عند مستوى ثقة **95%** .

الحل:

بما أن المجتمع محل الدراسة توزيعه الاحتمالي غير معروف، وحجم العينة كبير **n=100** ، فبتطبيق نظرية النهاية المركزية، حيث أن تباين المجتمع σ^2 معلوم ، فنحسب فترة الثقة المطلوبة من الفترة التالية :

$$\left(\bar{X} - Z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}} \quad , \quad \bar{X} + Z_{\alpha/2}\sqrt{\frac{\sigma^2}{n}}\right)$$

حيث:

$(1-\alpha)\ 100\% = 95\%$ ، $\bar{X} = 8525$ ، $n=100$ ، $\sigma = 3150$

$Z_{\alpha/2} = Z_{0.025} = 1.96$ ، $\frac{\alpha}{2} = 0.025$ ، $\alpha = 0.05$ ، $1-\alpha = 0.95$

وبالتعويض في الحد الأدنى والحد الأعلى، مع الانتباه أن القيمة التي عندنا هي قيمة الانحراف المعياري σ وليس σ^2، نحصل على:

الحد الأدنى لفترة الثقة:

$$\bar{X} - Z_{\propto/2}\sqrt{\frac{\sigma^2}{n}} = 8525 - (1.96)\sqrt{\frac{(3150)^2}{100}} = 8525 - 617.400 = 7907.600$$

الحد الأعلى لفترة الثقة:

$$\bar{X} + Z_{\propto/2}\sqrt{\frac{\sigma^2}{n}} = 8525 + (1.96)\sqrt{\frac{(3150)^2}{100}} = 8525 + 617.400 = 9142.400$$

إذن فترة الثقة للوسط الحسابي للمبالغ المودعة في كل الحسابات الجارية في هذا المصرف، عند مستوى ثقة **95%** هي:

(7907.600 ، 9142.400)

أي نستطيع القول بثقة قدرها **95%** بأن الوسط الحسابي للمبالغ المودعة في كل الحسابات الجارية في هذا المصرف، تقع بين القيمتين **7907.600** و **9142.400**.

(5-3-2) عندما يكون تباين المجتمع σ^2 مجهولا :

إذا كان لدينا مجتمع يتوزع توزيعا طبيعيا وكان تباينه σ^2 مجهولا ، فلكي نحصل على فترة الثقة للوسط الحسابي للمجتمع μ، نستخدم تباين العينة S^2 كمقدِّر بالقيمة للتباين المجهول σ^2 ، حيث تباين العينة يُحسب كما يلي :

$$S^2 = \frac{\sum(X - \bar{X})^2}{n - 1}$$

وإذا استعملنا قيمة S^2 بدلا من σ^2 سنحصل على المتغير العشوائي التالي :

$$T = \frac{\bar{X} - \mu}{\sqrt{\frac{S^2}{n}}}$$

ويتبع هذا المتغير توزيع t بدرجة حرية ***V=n-1***، وقد درسنا هذا التوزيع في الفصل السابق ، فقد علمنا أن توزيع t توزيع متماثل حول وسطه الحسابي الذي يساوى صفراً ، ويعنى ذلك أنه إذا كانت لدينا قيمتان للمتغير العشوائي **T** وكانت

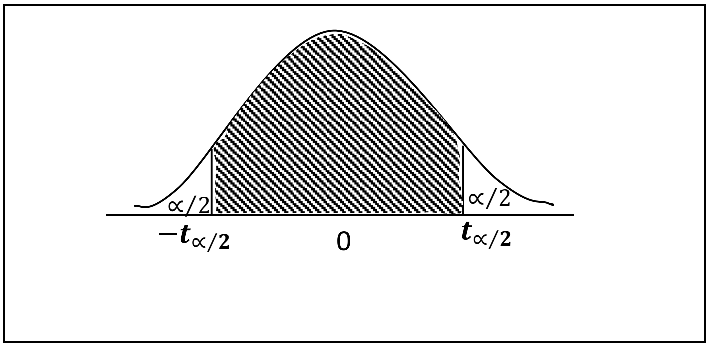
المساحة على يمين إحدى هاتين القيمتين تساوى المساحة على يسار القيمة الأخرى فستكون القيمتان متساويتين في القيمة المطلقة ومختلفتين في الإشارة .

وبصفة عامة إذا رمزنا للمساحة بين قيمتين للمتغير العشوائي **T** بالرمز **$1-\alpha$** بحيث تكون المساحة على يمين القيمة الكبرى تساوى المساحة التي على يسار القيمة الصغرى تساوى كلا منهما **$\alpha/2$**.

فستكون هاتان القيمتان متساويتين في القيمة المطلقة ومختلفتين في الإشارة، وسنرمز للقيمة الكبرى التي على يمينها المساحة $\alpha/2$ بالرمز $t\alpha_{/2}$، وبالتالي سنرمز للقيمة الأخرى التي على يسارها المساحة **$\alpha/2$ بالرمز** $-t\alpha_{/2}$، وذلك كما هو واضح في **شكل (5-3)**.

وبما أن المساحة تحت أي منحنى احتمالي هي عبارة عن احتمالات، فيعنى ذلك **أن** احتمال أن يأخذ المتغير العشوائي **T** عند درجة حرية **V** قيمة محصورة بين القيمتين $-\mathbf{t}\alpha_{/2}$ و $\mathbf{t}\alpha_{/2}$ يساوى **$1-\alpha$** ونعبر عن ذلك كما يلي :

$$P\left(-\mathbf{t}\alpha_{/2}(\nu) \leq \mathbf{T} \leq \mathbf{t}\alpha_{/2}(\nu)\right) = \mathbf{1}-\boldsymbol{\alpha}$$



شكل (5-3)

بما أن:
$$T = \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}}$$

إذن نستطيع كتابة الاحتمال السابق كما يلي:

$$P\left(-t_{\alpha/2}(\nu) \le \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}} \le t_{\alpha/2}(\nu)\right) = 1-\alpha$$

بإجراء بعض العمليات الجبرية على الحدود الثلاثة للمتباينة الموجودة بين القوسين، سنحصل على الصورة التالية للاحتمال:

$$P\left(\bar{X} - t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}} \le \mu \le \bar{X} + t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}}\right) = 1-\alpha$$

وبالتالي فإن فترة الثقة للوسط الحسابي للمجتمع **μ** عندما يكون تباين المجتمع مجهولا وعند مستوى ثقة **100%(1−α)** هي :

$$\left(\bar{X} - t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}} \quad ، \quad \bar{X} + t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}}\right)$$

مثال (5-4) :

إذا علمت أنه خلال فترة معينة كانت أوزان أكياس المكرونة المصنعة في أحد المصانع تتوزع توزيع طبيعيا، فإذا سحبنا من هذا المجتمع عينة تشمل **25** كيس مكرونة، ووجدنا أن الوسط الحسابي لأوزان الأكياس المسحوبة **498.25** جرام بتباين **1.69**، فقدّر متوسط أوزان كل أكياس المكرونة المصنعة في هذا المصنع خلال هذه الفترة وذلك باستخدام فترة ثقة عند مستوى ثقة **90%**.

الحل: يتكون المجتمع في هذه الدراسة من أوزان كل الأكياس المصنعة في هذا المصنع خلال هذه الفترة والتي تتبع توزيعا طبيعيا بتباين مجهول، ولذلك سنستعمل فترة الثقة التالية:

$$\left(\bar{X} - t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}} \quad ، \quad \bar{X} + t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}}\right)$$

حيث:

$(1 - \alpha)\ 100\% = 90\%$ ، $S^2 = 1.69$ ، $\bar{X} = 498.25$ ، $n=25$

إذن:

$\alpha/2 = 0.05$ ، $\alpha = 0.10$ ، $1 - \alpha = 0.90$

$t_{\alpha/2}(\nu) = t_{0.05}(24) = 1.711$ ، $v=n-1=25-1=24$

وبالتعويض في الحد الأدنى والحد الأعلى لفترة الثقة نحصل على:

الحد الأدنى لفترة الثقة:

$$\bar{X} - t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}} = 498.25 - (1.711)\sqrt{\frac{1.69}{25}} = 498.25 - 0.44 = 497.81$$

الحد الأعلى لفترة الثقة:

$$\bar{X} + t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}} = 498.25 + (1.711)\sqrt{\frac{1.69}{25}} = 498.25 + 0.44 = 498.69$$

إذن فترة الثقة لمتوسط أوزان كل الأكياس عند مستوى ثقة **90%** هي :

(497.81 ، 498.69)

أي نستطيع القول بثقة قدرها **90%** بأن الوسط الحسابي الحقيقي لأوزان كل الأكياس المصنعة في هذا مصنع يقع بين القيمتين **497.81** جرام و **498.69** جرام.

مثال (5-5) :

إذا علمت أن قيمة المبيعات اليومية لأحد المحلات التجارية تتبع توزيعا طبيعيا، والبيانات التالية تبين قيمة المبيعات اليومية خلال **8** أيام، اختيرت عشوائيا في سنة معينة:

173.00 ، 124.25 ، 130.00 ، 180.75

84.00 ، 132.50 ، 99.00 ، 140.50

باستخدام هذه البيانات قدّر الوسط الحسابي لكل المبيعات اليومية لهذا المحل في هذه السنة وذلك باستخدام فترة ثقة **95%**.

الحل:

يتكون المجتمع في هذه الدراسة من قيم كل المبيعات اليومية خلال تلك السنة المعنية بالدراسة، وحيث أن هذه القيم (المجتمع) تتبع توزيعا طبيعيا بتباين مجهول، لذلك سنستعمل فترة الثقة التالية:

$$\left(\overline{X} - t\alpha_{/2}(\nu)\sqrt{\frac{S^2}{n}} \quad ، \quad \overline{X} + t\alpha_{/2}(\nu)\sqrt{\frac{S^2}{n}} \right)$$

حيث:

(1-α) 100% = 95% ، n=8

إذن:

$\frac{\alpha}{2} = 0.025$ ، $\alpha = 0.05$ ، $1-\alpha = 0.95$

$t\alpha_{/2}(v) = t_{0.025}(7) = 2.365$ ، $v = n - 1 = 8-1 = 7$

أما قيمة الوسط الحسابي للعينة وتباين العينة يجب حسابهما من البيانات المعطاة، وذلك كما يلي:

$$\overline{X} = \frac{\sum_{i=1}^{n} Xi}{n} = \frac{180.75 + 130.00 + \cdots + 84.00}{8} = 133$$

$$S^2 = \frac{\sum(X - \overline{X})^2}{n-1}$$

$$= \frac{(180.75 - 133)^2 + (130.00 - 133)^2 + \cdots + (84.00 - 133)^2}{8-1} = 1082.73$$

وبالتعويض في الحد الأدنى والحد الأعلى لفترة الثقة نحصل على:

الحد الأدنى لفترة الثقة يساوي:

$$\overline{X} - t\alpha_{/2}(\nu)\sqrt{\frac{S^2}{n}} = 133 - (2.365)\sqrt{\frac{1082.73}{8}} = 133 - 27.51 = 105.49$$

الحد الأعلى لفترة الثقة تساوي:

$$\bar{\mathbf{X}} + t_{\alpha/2}(\nu)\sqrt{\frac{S^2}{n}} = 133 + (2.365)\sqrt{\frac{1082.73}{8}} = 133 + 27.51 = 160.51$$

إذن فترة الثقة للوسط الحسابي لكل المبيعات اليومية لهذا المحل خلال السنة المعنية بالدراسة، وعند مستوى ثقة **95% هي:**

(105.49 ، 160.51)

أي نستطيع القول بثقة قدرها **95%** بأن الوسط الحسابي لكل المبيعات اليومية لهذا المحل في تلك السنة يقع بين القيمتين **105.49** دينار و **160.51** دينار.

ملخص الفصل الخامس

التقدير هو أحد فرعى الاستنتاج الإحصائي، فهو يهتم بكيفية تقدير معالم مجتمع باستخدام بيانات نحصل عليها من عينة عشوائية مسحوبة من هذا المجتمع، ويوجد نوعان من التقدير:

1. التقدير بقيمة:

هو تقدير معلمة المجتمع بإحصاءة نحسب قيمتها من العينة، أي نقدر المعلمة المجهولة بقيمة واحدة فقط، وأفضل مقدِّر بالقيمة للوسط الحسابي للمجتمع μ هو الوسط الحسابي للعينة $\overline{X}$ وأفضل مقدِّر بالقيمة لتباين المجتمع σ^2 هو تباين العينة S^2.

2. التقدير بفترة:

المقصود به استخدام التقدير بقيمة لإنشاء فترة نعتقد وقوع المعلمة المجهولة بداخلها بدرجة ثقة معينة، فوجدنا أنه عندما يتبع المجتمع التوزيع الطبيعي أو حجم العينة كبير، فإن فترة الثقة للوسط الحسابي للمجتمع عندما يكون تباين المجتمع معلوما، وعند مستوى ثقة **(1−α) 100%** هي:

$$\left(\overline{\mathbf{X}} - \mathbf{Z}\alpha_{/2}\sqrt{\sigma^2/n} \quad ، \quad \overline{\mathbf{X}} + \mathbf{Z}\alpha_{/2}\sqrt{\sigma^2/n}\right)$$

وفترة الثقة للوسط الحسابي للمجتمع عندما يكون تباين المجتمع مجهولا، وعند مستوى ثقة **(1−α) 100%**:

$$\left(\overline{\mathbf{X}} - \mathbf{t}\alpha_{/2}(\nu)\sqrt{S^2/n} \quad ، \quad \overline{\mathbf{X}} + \mathbf{t}\alpha_{/2}(\nu)\sqrt{S^2/n}\right)$$

مع العلم أن: طول أية فترة = الحد الأعلى – الحد الأدنى.

تمارين (5)

1. ما المقصود بالتقدير بقيمة؟
2. ما المقصود بالتقدير بفترة؟
3. إذا كان مجتمع وسطه الحسابي مجهولاً، وسحبنا منه العينة العشوائية التاليـــــــة: **15**، **7** ، **16** ، **18** ، **14** ، **8** ، قدر بقيمة الوسط الحسابي لهذا المجتمع .
4. إذا كان الدخل الشهري للعائلات القاطنة في مدينة ما يتبع توزيعا طبيعيا بوسط حسابي مجهول وتباين يساوى **400**، فإذا اخترنا من هذه المدينة عينة عشوائية تشمل **25** عائلة، ووجدنا أن الوسط الحسابي للدخل الشهري لهذه العينة يساوي **190** دينار. فقدّر الوسط الحسابي للدخل الشهري لكل العائلات القاطنة في هذه المدينة وذلك:
   - **i** : بالتقدير بقيمة .
   - **ii** : باستخدام فترة ثقة عند مستوى ثقة **95%** .
5. إذا كان وزن إنتاج شجيرات الطماطم في إحدى المزارع يتبع توزيعا طبيعيا بوسط حسابي مجهول وتباين =**4**، اختيرت عينة عشوائية تحتوي على **36** شجيرة، فوجد أن الوسط الحسابي لوزن إنتاجها **13.5** كيلو جرام. فقدّر الوسط الحسابي لوزن إنتاج شجيرات كل المزرعة وذلك باستخدام فترة ثقة عند مستوى ثقة **90%**.
6. إذا كان المبلغ الشهري لفاتورة الكهرباء في مدينة معينة يتبع توزيعا ما، تباينه يساوى **16**، فإذا اخترنا من هذه المدينة عينة عشوائية تشمل **100** عائلة، ووجدنا أن الوسط الحسابي للمبلغ الشهري لفاتورة الكهرباء لهذه العينة يساوى **75** دينار. فقدّر الوسط الحسابي للمبلغ الشهري لفاتورة الكهرباء لكل العائلات القاطنة في هذه المدينة، وذلك باستخدام فترة ثقة عند مستوى ثقة **95%**.

7. إذا كان طول الأنابيب المنتجة من مصنع للأنابيب يتبع توزيعا طبيعيا، بوسط حسابي وتباين مجهولين، واخترنا عينة عشوائية تحتوى على 5 أنابيب وكانت أطوالها كما يلي: **11 ، 13 ، 21 ، 9 ، 15** .

فقدّر الوسط الحسابي لأطوال كل الأنابيب المصنوعة في هذا المصنع وذلك:

i: بالتقدير بقيمة .

ii: باستخدام فترة ثقة عند مستوى ثقة 99%.

8. إذا كان عدد السيارات التي تبيعها وكالة من وكالات السيارات أسبوعيا يتبع توزيعا طبيعيا، بوسط حسابي وتباين مجهولين، فإذا اخترنا عينة عشوائية تشمل المبيعات الأسبوعية لهذه الوكالة لمدة **10** أسابيع خلال سنة معينة، وكانت البيانات كما يلي:

2 ، 1 ، 0 ، 3 ، 4 ، 1 ، 1 ، 3 ، 2 ، 3

قدّر الوسط الحسابي لعدد السيارات المباعة أسبوعيا من قبل هذه الوكالة خلال تلك السنة، وذلك باستخدام فترة ثقة،

i: عند مستوى ثقة 90% .

ii: عند مستوى ثقة 95% .

الفصل السادس
اختبارات الفروض

لقد ذكرنا أن مادة الإحصاء الاستدلالي تنقسم إلى نوعين، هما التقدير واختبارات الفروض، وقد عرضنا في الفصل السابق موضوع التقدير، وفي هذا الفصل سنعرض موضوع اختبارات الفروض الإحصائية.

(6–1) تعريف الفروض الإحصائية:

الفرضية الإحصائية هي عبارة عن تخمينات أو تعبيرات حول معلمة مجهولة من معالم المجتمع الإحصائي، وقد سميت بالفروض لأنها قد تكون صحيحة أو غير صحيحة.

وسوف نتطرق فيما يلي إلى أنواع الفروض الإحصائية وخطوات اختبارها.

(6–2) أنواع الفروض الإحصائية:

يوجد نوعان من الفروض الإحصائية هما: فرض العدم والفرض البديل، ويمكن تعريفهما كما يلي:

1. فرض العدم:

فرض العدم هو التخمين أو التعبير الذي يأمل الباحث الإحصائي أن يرفضه، وهو الفرض الذي يعطى للمعلمة قيمة يعتقد الباحث أنها ليست القيمة الحقيقية للمعلمة، ولذلك قام بإجراء الاختبار، ومن هنا جاءت تسميته بفرض العدم، أي عدم تمثيل التعبير المذكور في هذا الفرض للقيمة الحقيقية للمعلمة، وفرض العدم هو الفرض الذي يحتوى على إشارة المساواة فهو يدل عادة على عدم الاختلاف، فإذا كان الاختبار خاصا بمعلمة واحدة فيفترض الفرض مساواة هذه المعلمة لقيمة معينة،

وإذا كان الفرض خاصا بمقارنة معلمتين في مجتمعين فيفترض فرض العدم تساوى هاتين المعلمتين وهكذا ..... ويرمز لفرض العدم بالرمز H_0 .

2. الفرض البديل:

هو الفرض الذي يُقبل كبديل لفرض العدم عند رفض فرض العدم، ويرمز له بالرمز H_1 .

فإذا كان الوسط الحسابي لمجتمع معين μ غير معروف ، ونريد أن نختبر أن قيمة هذا الوسط الحسابي تساوى قيمة معينة ولتكن μ_0 أم لا ؟ فتكتب الفروض في هذه الحالة كما يلي :

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

وإذا كان فرض العدم يعتبر أن الوسط الحسابي المجهول μ أقل من أو يساوى قيمة معينة μ_0، ففي هذه الحالة تكتب الفروض الإحصائية كما يلي :

$$H_0: \mu \leq \mu_0$$
$$H_1: \mu > \mu_0$$

ونستطيع في هذه الحالة أن نكتب في فرض العدم إشارة المساواة فقط أي $H_0: \mu = \mu_0$ وحيث أن الفرض البديل يحتوي على إشارة أكبر من فقط، فنفهم ضمنيا أن إشارة أقل من، يجب أن تكون في فرض العدم حتى إذا لم تذكر صراحة.

أما إذا كان فرض العدم يعتبر أن الوسط الحسابي المجهول μ أكبر من أو يساوى قيمة معينة ، ففي هذه الحالة نكتب الفروض الإحصائية كما يلي :

$$H_0: \mu \geq \mu_0$$
$$H_1: \mu < \mu_0$$

وفي هذه الحالة أيضا نستطيع أن نكتب في فرض العدم إشارة المساواة فقط، أي $H_0: \mu = \mu_0$ وحيث أن الفرض البديل يحتوي على إشارة (أقل من) فقط فنفهم ضمنيا أن إشارة (أكبر من) يجب أن تكون في فرض العدم حتى إذا لم تذكر صراحة.

ويعتمد رفض فرض العدم أو عدم رفضه على أساس قاعدة يضعها متخذو القرارات استنادا على خبرتهم السابقة، وعلى أساس البيانات المتوفرة من العينة المسحوبة.

فإذا كانت نتائج العينة تؤيد فرض العدم وفقا للقاعدة الموضوعة، فإننا لا نرفض فرض العدم (نقبله) لعدم وجود دليل كاف لرفضه، وإذا كانت بيانات العينة لا تؤيد فرض العدم وفقا للقاعدة الموضوعة فإننا نستطيع أن نرفض فرض العدم.

ولاتخاذ القرار برفض أو قبول فرض العدم H_0 نعتمد على ما يسمى بإحصاءة الاختبار حيث تعرف كما يلي :

(3-6) تعريف إحصاءة الاختبار:

هي متغير عشوائي يجب أن يكون توزيعه الاحتمالي معلوما عندما يكون فرض العدم H_0 صحيحا، ونحسب قيمتها من بيانات العينة، وتستخدم قيمة إحصاءة الاختبار المحسوبة من بيانات العينة العشوائية المسحوبة من المجتمع محل الدراسة والتي يطلق عليها القيمة المشاهدة لإحصاءة الاختبار لاتخاذ القرار برفض أو قبول فرض العدم H_0 .

ويتم تقسيم كل القيم التي يمكن أن تأخذها إحصاءة الاختبار، لمجموعتين غير متداخلتين، أحدهما للنتائج التي إذا ظهرت نقبل فرض العدم وتسمى منطقة القبول، والأخرى للنتائج التي إذا ظهرت نرفض فرض العدم وتسمى منطقة الرفض. أي أن يقسم توزيع المعاينة لإحصاءة الاختبار إلى منطقتين:

1. **منطقة القبول:** هي المنطقة التي تحتوى على قيم إحصاءة الاختبار التي تؤدى إلى قبول فرض العدم H_0 .

2. **منطقة الرفض:** هي المنطقة التي تحتوى على قيم إحصاءة الاختبار التي تؤدى إلى رفض فرض العدم H_0 .

والقيمة التي تفصل بين هاتين المنطقتين تسمى **بالقيمة الحرجة.**

يكون القرار رفض فرض العدم إذا كانت القيمة المشاهدة لإحصاءة الاختبار تقع في منطقة الرفض، وقبول فرض العدم إذا وقعت القيمة المشاهدة لإحصاءة الاختبار في منطقة القبول.

(6-4) أنواع الأخطاء :

حيث أن اتخاذ القرار يعتمد على القيمة المشاهدة لإحصاءة الاختبار، أي على القيمة المحسوبة من العينة المختارة، وقد تكون هذه العينة لا تمثل المجتمع الذي سحبت منه تمثيلا صحيحًا، مما يؤدى إلى وقوع متخذ القرار في خطأ من اثنين:

1. **خطأ من النوع الأول:**

يحدث هذا الخطأ إذا كان فرض العدم، في الحقيقة صحيحا، ولكن بيانات العينة تظهر أنه غير صحيح، أي أن نتائج العينة تؤدى إلى رفض فرض العدم مع أنه في الواقع صحيح. ويرمز لاحتمال وقوع خطأ من النوع الأول، أي لاحتمال رفض فرض العدم مع أنه في الواقع صحيح بالرمز α ويطلق عليه **مستوى المعنوية**، أي أن:

$$\alpha = P(\text{ارتكاب خطأ من النوع الأول})$$

$$= P(\text{رفض فرض العدم} \,/\, \text{فرض العدم صحيح})$$

2. خطأ من النوع الثاني:

يحدث هذا الخطأ نتيجة لقبول فرض العدم مع أنه في الواقع غير صحيح، أي أن بيانات العينة تؤيد فرض العدم مع أن فرض العدم في الحقيقة غير صحيح، ويـــرمز إلى احتمال وقوع خطأ من النوع الثاني، أي احتمال قبول فرض العدم مع أن فرض العدم في الواقع خطأ بالرمز **β** أي أن :

β = P (ارتكاب خطأ من النوع الثاني)

= P (فرض العدم خطأ / رفض فرض العدم)

ويمكن تلخيص الحالات التي يتعـرض لها متخـذ القـــرار في الجـــدول التالـــي:

جدول (6–1)

القرار	H_0 صحيح	H_0 غير صحيح
قبول H_0	قرار سليـم	خطأ من النوع الثاني
رفض H_0	خطأ من النوع الأول	قرار سليم

وعادة تحدد قيمة **α** بالقيمة **0.10** أو **0.05** **أو** **0.01** والاختيار بين هذه القيم يعتمد على الاعتقاد الشخصي لمتخذ القرار ومدى خبرته، وبالطبع كلما زادت خطورة رفض فرض العدم كلما قلت قيمة **α** المستعملة .

وحيث أن **α** تمثل احتمال رفض فرض العدم مع صحته، ونعلم أن أي احتمال هو مساحة، وبالتالي فإن ∝ تمثل مساحة منطقة الرفض، فبمعرفة **α** نستطيع تحديد منطقة الرفض على الشكل الذي يمثل توزيع المعاينة لإحصاءة الاختبار عندما يكون

H_0 صحيحا، وحيث أن المساحة الكلية تحت منحنى دالة كثافة الاحتمال تساوى الواحد الصحيح فنستطيع تحديد منطقة القبول ، وتكون هي المنطقة تحت المنحنى المــكملة لمنطقة الرفض ومساحتها **(1-α)**. وعند تحديد منطقة الرفض سنتعرض إلى الحالات الثلاثة التالية:

1. تكون منطقة الرفض موزعة على طرفي التوزيع الاحتمالي لإحصاءة الاختبار إذا كانت الفروض الإحصائية كما يلي:

$$H_0: \mu = \mu_0$$
$$H_1: \mu \neq \mu_0$$

ويسمى الاختبار في هذه الحالة **اختبار ذو طرفين.**

2. تكون منطقة الرفض كلها في الطرف الأيمن للتوزيع الاحتمالي لإحصاءة الاختبار إذا كانت الفروض الإحصائية كما يلي:

$$H_0: \mu = \mu_0$$
$$H_1: \mu > \mu_0$$

ويسمى الاختبار في هذه الحالة **اختبار ذو طرف واحد أيمن.**

3. تكون منطقة الرفض كلها في الطرف الأيسر للتوزيع الاحتمالي لإحصاءة الاختبار إذا كانت الفروض الإحصائية كما يلي:

$$H_0: \mu = \mu_0$$
$$H_1: \mu < \mu_0$$

ويسمى الاختبار في هذه الحالة **اختبار ذو طرف واحد أيسر.**

وبعد تحديد منطقة الرفض ومنطقة القبول على الشكل الذي يمثل التوزيع الاحتمالي لإحصاءة الاختبار، نحسب القيمة المشاهدة لإحصاءة الاختبار، وهي قيمة إحصاءة الاختبار المحسوبة من واقع بيانات العينة، فإذا وقعت القيمة المشاهدة لإحصاءة الاختبار في منطقة القبول نقبل فرض العدم $\mathbf{H_0}$، وإذا وقعت القيمة المشاهدة لإحصاءة الاختبار في منطقة الرفض نرفض فرض العدم $\mathbf{H_0}$

باستخدام مستوى معنوية α، فيجب ذكر مستوى المعنوية عند اتخاذ القرار ، وذلك لأن القرار قد يختلف باختلاف مستوى المعنوية المستخدم .

ويمكن تلخيص خطوات اختبارات الفروض الإحصائية فيما يلي:

(6-5) خطوات اختبارات الفروض الإحصائية:

1. صياغة فرض العدم والفرض البديل.
2. تحديد مستوى المعنوية α **(مساحة منطقة الرفض)** .
3. اختيار إحصاءة الاختبار المناسبة وهي الإحصاءة التي تعتمد على أفضل مقدِّر بالقيمة للمعلمة المجهولة التي نجرى الاختبار بخصوصها ويجب معرفة التوزيع الاحتمالي لهذه الإحصاءة عندما يكون H_0 صحيحا، وذلك لتحديد منطقة الرفض ومنطقة القبول .
4. حساب القيمة المشاهدة لإحصاءة الاختبار، أي حساب قيمة إحصاءة الاختبار من واقع البيانات المشاهدة التي نحصل عليها من العينة وذلك مع افتراض صحة فرض العدم.
5. اتخاذ القرار المناسب، ويكون أحد القرارين التاليين:

أ. نرفض فرض العدم H_0 إذا وقعت القيمة المشاهدة لإحصاءة الاختبار في منطقة الرفض.

ب. نقبل فرض العدم H_0 إذا وقعت القيمة المشاهدة لإحصاءة الاختبار في منطقة القبول.

(6–6) اختبارات للوسط الحسابي للمجتمـــــع μ:
(1–6–6) عندما يكون تباين المجتمع σ^2 معلوما:

نعلم أن أفضل مقدّر بالقيمة للوسط الحسابي للمجتمع المجهول **μ** هو الوسط الحسابي للعينة $\bar{X}$، وأن توزيع المعاينة للوسط الحسابي للعينة $\bar{X}$ توزيع طبيعي إذا كان المجتمع المسحوبة منه العينة يتوزع توزيعا طبيعيا، وأن توزيع المعاينة للوسط الحسابي للعينة قريبا من التوزيع الطبيعي إذ كان حجم العينة أكبر من ثلاثين **(*n*>30)**، وذلك بغض النظر عن توزيع المجتمع المسحوبة منه العينة (وفقا لنظرية النهاية المركزية) . وحيث أننا عند التعامل مع أي متغير طبيعي يجب تحويله إلى صيغته المعيارية، فتكون إحصاءة الاختبار المناسبة لإجراء اختبار الوسط الحسابي للمجتمع عندما يكون σ^2 معلوما ، هي المتغير العشوائي الطبيعي المعياري Z حيث

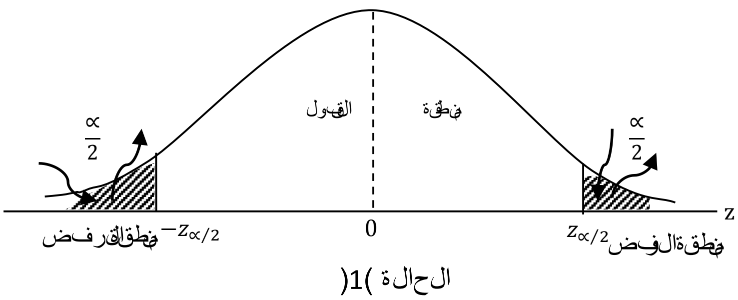
$$\boxed{Z=\frac{\bar{X}-\mu}{\sqrt{\sigma^2/n}}}$$

وعند استخدام مستوى معنوية يساوى **α** فيعنى ذلك أن مساحة منطقة الرفض تحت منحنى التوزيع الطبيعي المعياري تساوى **α** وسنتعرض لإحدى الحالات الثلاثة التالية:

1. إذا كان الاختبار ذا طرفين فستكون مساحة منطقة الرفض **α** مقسومة إلى منطقتين منطقة في الطرف الأيمن لتوزيع المعاينة لإحصاءة الاختبار (**التوزيع الطبيعي المعياري**) ومنطقة أخرى مساوية لها في الطرف الأيسر للتوزيع ، وبالتالي ستكون مساحة كل منطقة تساوى $\alpha/_2$ وتكون المساحة بين هاتين المنطقتين هي منطقة القبول . ونحدد القيم الحرجة التي تفصل منطقة القبول عن منطقتي الرفض من جدول (م.1) فتكون القيمة الحرجة التي على اليمين هي قيمة Z الجدولية التي على يمينها مساحة قدرها $\alpha/_2$ ونرمز لها $Z_{\alpha/2}$ وتكون القيمة الحرجة التي على اليسار

هي القيمة الجدولية التي على يسارها مساحة قدرها $\alpha/_2$، وبما أن المنحنى الطبيعي المعياري متماثل فستكون هي نفسها القيمة الحرجة التي على اليمين مع اختلاف الإشارة، وبالتالي نرمز لها بالرمز $-Z\alpha/_2$، **الشكل (6-1)** يوضح منطقتي الرفض القبول في هذه الحالة ونرفض فرض العدم $\mathbf{H_0}$ إذا كانت القيمة المشاهدة لإحصاءة الاختبار تقع في منطقة الرفض، أي إذا كانت القيمة المشاهدة لإحصاءة الاختبار (القيمة المحسوبة **Z**) أكبر من القيمة الجدولية الموجبة أو أقل من القيمة الجدولية السالبة، أي:

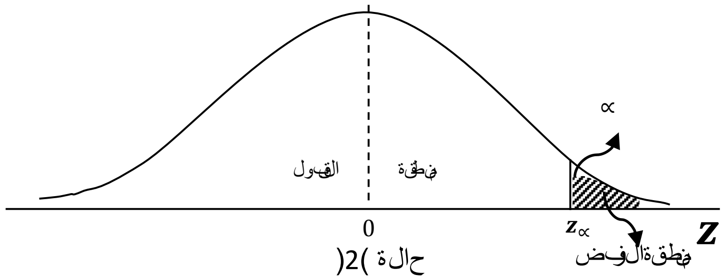
$$z \leq -z\propto/_2 \quad و \quad z \geq z\propto/_2$$



شكل (6-1) اختبار ذو طرفين

2. أما إذا كان الاختبار ذو طرف أيمن، فستكون منطقة الرفض كلها ناحية اليمين ومساحتها تساوى $\boldsymbol{\alpha}$، وتكون القيمة الحرجة التي تفصل بين منطقة الرفض ومنطقة القبول هي القيمة الجدولية التي على يمينها مساحة قدرها $\boldsymbol{\alpha}$، ويرمز لها بالرمز Z_α، وشكل **(6-2)** يوضح ذلك، ونرفض فرض العدم H_0 إذا كانت القيمة المشاهدة لإحصاءة الاختبار أكبر من القيمة الجدولية، أي:

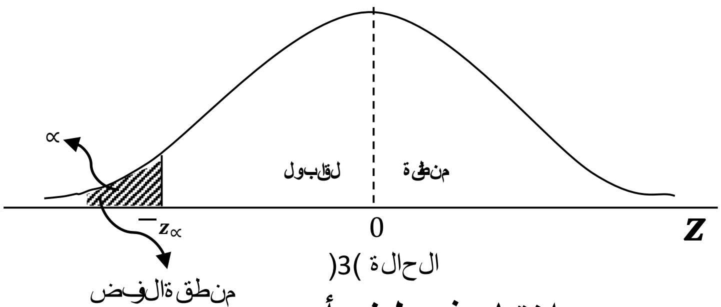
$$Z \geq Z_{\propto}$$



شكل (6−2) اختبار ذو طرف أيمن

3. أما إذا كان الاختبار ذو طرف أيسر، فستكون منطقة الرفض كلها ناحية اليسار ومساحتها تساوى **α**، وتكون القيمة الحرجة التي تفصل بين منطقة الرفض ومنطقة القبول هي القيمة الجدولية التي على يسارها مساحة قدرها **α**، ويرمز لها بالرمز $-Z_{\propto}$ وشكل **(6−3)** يوضح ذلك ، ونرفض فرض العدم H_0 إذا كانت: القيمة المشاهدة لإحصاءة الاختبار أقل من القيمة الجدولية أي أن:

$$Z \leq -Z_{\alpha}$$



اختبار ذو طرف أيسر

شكل (6−3)

مثال (6-1) :

يدِّعى مدير مصنع لصناعة الأنابيب إن صناعة هذه الأنابيب في مصنعه دقيقة جدا ومطابقة للمواصفات وأن الوسط الحسابي لأطوال كل الأنابيب المنتجة **60** سم بتباين **2.25**، وللتأكد من صحة قوله سحبت عينه عشوائية من الإنتاج الكلى للمصنع تحتوي على **16** أنبوبًا، فكان الوسط الحسابي لأطوالها **58.8** سم، فإذا كانت أطوال الأنابيب تتبع توزيعا طبيعيا، فاختبر صحة ادعاء مدير المصنع باستخدام مستوى معنوية **0.05** .

الحل :

المجتمع في هذه الدراسة يتكون من أطوال كل الأنابيب المنتجة في هذا المصنع، وعند وضع الفروض، سيفترض الفرض H_0 أن إدعاء المدير صحيح والإنتاج في المصنع مطابق للمواصفات، أي أن الوسط الحسابي لأطوال كل الأنابيب المنتجة في المصنع μ يساوى **60** سم، أما الفرض البديل فسيفترض الحالة البديلة وهي أن الادعاء غير صحيح ، أي أن الإنتاج في المصنع غير مطابق للمواصفات، فقد يكون الوسط الحسابي لأطوال كل الأنابيب المنتجة في المصنع μ أقل أو أكبر من **60** سم وتكتب هذه الفروض كما يلي :

$$H_0: \mu = 60$$
$$H_1: \mu \neq 60$$

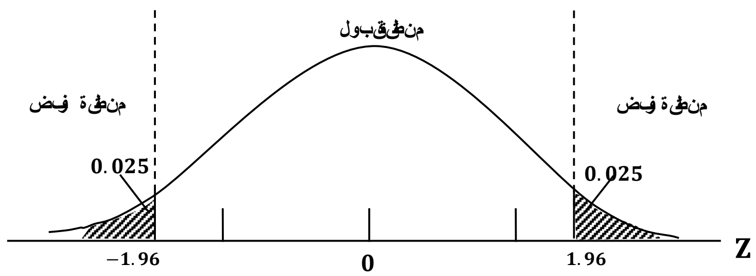
وبما أن المجتمع يتوزع توزيعا طبيعيا، وتباينه معلوم ($\sigma^2 = 2.25$)، إذن إحصاءة الاختبار المناسبة لهذه الحالة هي الإحصاءة التالية :

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$$

وتوزيع المعاينة لهذه الإحصاءة توزيع طبيعي معياري.

وحيث أن الاختبار ذو طرفين ومستوى المعنوية **0.05** = α فمن جدول (م.**1**) نستطيع الحصول على القيمة الحرجة للطرف الأيمن وهي:

$\mathbf{Z_{\alpha/2} = Z_{0.025} = 1.96}$ ، وبالتالي تكون القيمة الحرجة للطرف الأيسر تساوى **1.96** − والمساحة بين هاتين القيمتين تمثل منطقة القبول ، أما المساحة على الطرفين فتمثل منطقة الرفض ، وذلك كما هو واضح في الشكل (**6−4**) .



شكل (6−4)

وبعد تحديد منطقة القبول والرفض نقوم بحساب القيمة المشاهدة لإحصاءة الاختبار وهي قيمة الاحصاءة بعد أن نعوِّض فيها ببيانات العينة، فنجد أن القيمة المشاهدة لإحصاءة الاختبار هي:

$$Z = \frac{\overline{\mathbf{X}} - \boldsymbol{\mu}}{\sqrt{\frac{\boldsymbol{\sigma}^2}{\mathbf{n}}}} = \frac{\mathbf{58.8 - 60}}{\sqrt{\frac{\mathbf{2.25}}{\mathbf{16}}}} = -\frac{\mathbf{1.2}}{\mathbf{0.375}} = \mathbf{-3.2}$$

وبما أن القيمة المشاهدة لإحصاءة الاختبار تقع في منطقة الرفض، إذن نرفض H_0 بمستوى معنوية **0.05** ويعنى ذلك أن إدعاء مدير المصنع غير صحيح ، أي إن الوسط الحسابي لأطوال كل الأنابيب المنتجة في المصنع لا يساوى **60** سم. وبالطبع هذا القرار الذي اتخذناه ليس صحيحا **100%**، بل نتوقع ارتكاب خطأ، لأننا اعتمدنا في قرارنا على بيانات عينة وقد تكون هذه العينة لا تمثل المجتمع تمثيلا سليما،

واحتمال أن يكون هذا القرار خاطئ يساوى مستوى المعنوية = **0.05** (احتمال وقوع خطأ من النوع الأول).

مثال (6-2) :

إذا علمت من دراسة إحصائية سابقة أن أطوال كل طلبة مدرسة ثانوية ما تتبع توزيعا طبيعيا بوسط حسابي **165 سم** وتباين قدره **9**، وفي السنوات الأخيرة يعتقد أن الوسط الحسابي لأطوال كل طلبة هذه المدرسة قد زاد، ولذلك اختيرت عينة عشوائية تشمل **16** طالبا من طلبة هذه المدرسة، ووجد أن الوسط الحسابي لهذه العينة يساوى **165.75 سم**، فأختبر هذا الاعتقاد، وذلك باستخدام مستوى معنوية $\alpha = 0.05$.

الحل:

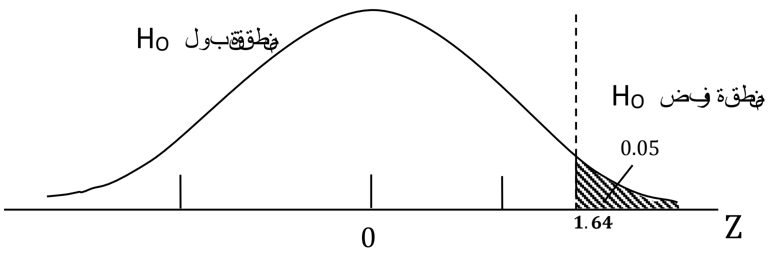
نكتب فروض هذا الاختبار كما يلي:

$$H_0: \mu = 165$$
$$H_1: \mu > 165$$

وبما أن الأطوال تتوزع توزيعا طبيعيا وتباين المجتمع معلوم فإن إحصاءة الاختبار المناسبة هي:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$$

وتوزيع المعاينة لهذه الإحصاءة توزيع طبيعي معياري. وحيث أن الاختبار ذو طرف أيمن ومستوى المعنوية **0.05** α فمن جدول **(م.1)** نستطيع الحصول على القيمة الحرجة $Z_a = Z_{0.05} = 1.64$، وتكون المساحة على يمين هذه القيمة تمثل منطقة الرفض والمساحة على يسارها تمثل منطقة القبول، وذلك كما هو واضح في الشكل **(6–5)** .



شكل (5−6)

ثم نحسب القيمة المشاهدة لإحصاءة الاختبار حيث:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} = \frac{165.75 - 165}{\sqrt{9/16}} = \frac{(0.75)(4)}{3} = 1.0$$

وبما أن القيمة المشاهدة لإحصاءة الاختبار تقع في منطقة القبول **(1.0< 1.64)**، إذن سيكون قرارنا هو قبول H_0 بمستوى معنوية α يساوى **0.05**. أي أن الوسط الحسابي لأطوال كل طلبة المدرسة الثانوية لم يزد.

مثال (3-6) :

تدّعى شركة لإنتاج الدقيق المعبأ في أكياس، بأن الوسط الحسابي لأوزان كل الأكياس المنتجة يساوى **99** جرام، بتباين **4**، ولكن المسؤلين على الرقابة يشكُون في ذلك ويعتقدون أن الوسط الحسابي لوزن الأكياس المنتجة أقل من **99** جرام، ولذلك اختاروا من إنتاج هذه الشركة عينة عشوائية تحتوى **100** كيس، وكان الوسط الحسابي لهذه العينة **98** جرام، فهل تؤيد هذه العينة ادعاء الشركة؟ اختبر ذلك باستخدام مستوى معنوية **0.02**.

الحل:

تكتب فروض هذا الاختبار كما يلي:

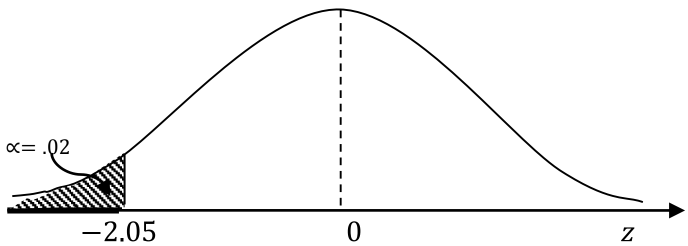
$$H_0: \mu = 99$$
$$H_1: \mu < 99$$

وبما أن حجم العينة كبير وتباين المجتمع معلوم فإن إحصاءة الاختبار المناسبة هي:

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$$

وتوزيع المعاينة لهذه الإحصاءة توزيع طبيعي معياري.

وحيث أن الاختبار ذو طرف أيسر ومستوى المعنوية $\alpha = 0.02$ فمن جدول (م.1) نستطيع الحصول على القيمة الحرجة $-Z_\alpha = -Z_{0.02} = -2.05$، وتكون المساحة على يسار هذه القيمة تمثل منطقة الرفض والمساحة على يمينها تمثل منطقة القبول وذلك كما هو واضح في الشكل (6-6) .



شكل (6-6)

ثم نحسب القيمة المشاهدة لإحصاءة الاختبار حيث :

$$Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}} = \frac{98 - 99}{\sqrt{4/100}} = \frac{-1.0}{\frac{2}{10}} = -5.0$$

وبما أن القيمة المشاهدة لإحصاءة الاختبار تقع في منطقة الرفض لأن ($-5.0 < -2.05$)، إذن فسيكون قرارنا هو رفض H_0 بمستوى معنوية $\alpha = 0.02$ ويعني ذلك أن الوسط الحسابي لوزن أكياس الدقيق المنتجة من هذه الشركة أقل من **99**، أي أن إدعاء الشركة غير صحيح .

(2-6-6) اختبارات للوسط الحسابي للمجتمع μ عندما يكون تباين المجتمع σ^2 مجهولا:

إذا كان تباين المجتمع σ^2 مجهولا فنقدره بأفضل مقدِّر له وهو تباين العينة S^2 حيث :

$$S^2 = \frac{\sum(X-\overline{X})^2}{n-1}$$

وتكون إحصاءة الاختبار المناسبة في هذه الحالة هي المتغير العشوائي **T** حيث :

$$T = \frac{\overline{X}-\mu}{\sqrt{S^2/n}}$$

ويتبع توزيع **t** بدرجات حرية **(V=n-1)**. وبالطبع في هذه الحالة نحصل على القيم الحرجة من جدول توزيع **t** ، جدول **(م.2)** .

مثال (6-4) :

في دراسة إحصائية سابقة وجد أن الوسط الحسابي للإنتاج السنوي للعامل في مصنع للسجاد **14** سجادة ، فإذا اتبع أسلوب جديد للإنتاج في هذا المصنع ، واخترنا عينة عشوائية تحتوى على **9** عاملين وكان إنتاجهم السنوي كما يلي :

18 ، 23 ، 17 ، 19 ، 15 ، 11 ، 14 ، 16 ، 11

بافتراض أن الإنتاج في المصنع يتبع توزيعا طبيعيا ، أختبر ما إذا كانت الطريقة الجديدة أدت إلى زيادة الوسط الحسابي للإنتاج السنوي لكل العاملين في هذا المصنع، وذلك باستخدام مستوى معنوية $\alpha = 0.05$.

الحل :

الفروض الإحصائية لهذا الاختبار كما يلي :

$$H_0: \mu = 14$$
$$H_1: \mu > 14$$

بما أن المجتمع يتوزع توزيعا طبيعياً وتباين المجتمع مجهولا فإن إحصاءة الاختبار المناسبة في هذه الحالة هي:

$$T = \frac{\overline{X}-\mu}{\sqrt{S^2/n}}$$

وتوزيعها الاحتمالي هو توزيع **t** بدرجات حرية **(V=n-1=9-1=8)**.

وحيث أن الاختبار ذو طرف أيمن ومستوى المعنوية $\boldsymbol{\alpha} = 0.05$ فمن جدول (م.2) نحصل على القيمة الحرجة $\mathbf{t}_{\alpha}(v) = t_{0.05}(8) = 1.86$ ، وتكون المساحة على يمين هذه القيمة تمثل منطقة الرفض والمساحة على يسارها تمثل منطقة القبول وذلك كما هو واضح في الشكل **(6−7)**.

ولكي نحسب القيمة المشاهدة لإحصاءة الاختبار، يجب أولا حساب الوسط الحسابي للعينة $\overline{X}$ وتباين العينة S^2 وذلك كما يلي :

$$\overline{X} = \frac{\Sigma X}{n} = \frac{11 + 16 + 14 + \cdots + 18}{9} = 16$$

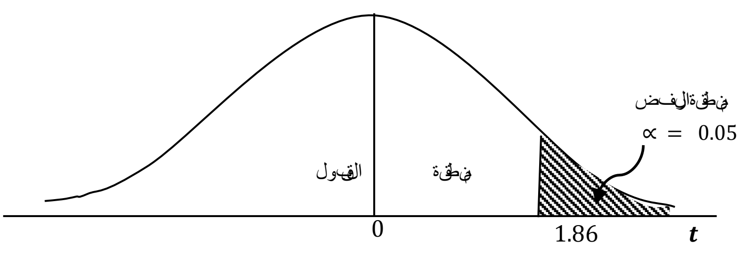
$$S^2 = \frac{\sum(X - \overline{X})^2}{n - 1}$$

$$= \frac{(11 - 16)^2 + (16 - 16)^2 + \cdots + (18 - 16)^2}{9 - 1} = \frac{118}{8} = 14.75$$

إذن :

$$t = \frac{\overline{X} - \mu}{\sqrt{S^2/n}} = \frac{16 - 14}{\sqrt{14.75/9}} = \frac{2}{1.28} = 1.56$$

وبمـا أن القيمـة المشاهـدة لإحصـاءة الاختبـار تقـع في منطقة القبـول لأن **($1.56 < 1.86$)**، إذن يكون قرارنا هو قبول H_0 بمستوى معنوية $\alpha = 0.05$ ، ويعنى ذلك أن الوسط الحسابي للإنتاج السنوي للعامل لم يزد، أي أن الطريقة الجديدة لم تؤدى إلى زيادة الوسط الحسابي للإنتاج السنوي للعامل .



شكل (6−7)

مثال (6-5):

عينة عشوائية تشمل **16** عائلة اختيرت من مدينة ما، وكان الوسط الحسابي لدخول هذه العائلات يساوى **205** دينار ، وتباينها يساوى **900**، فإذا علمت أن دخول كل العائلات في هذه المدينة تتبع توزيعا طبيعيا ، فباستخدام بيانات هذه العينة، أختبر ما إذا كان الوسط الحسابي لدخول كل العائلات في هذه المدينة أقل من **210** دينار، وذلك عند مستوى معنوية يساوى **0.10** .

الحل:

فروض الاختبار المطلوب كما يلي:

$$H_0: \mu = 210$$
$$H_1: \mu < 210$$

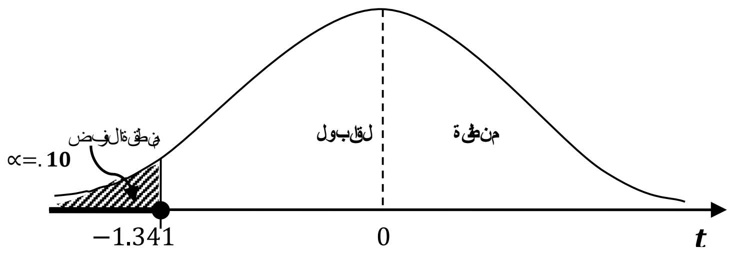
بما أن المجتمع يتوزع توزيعا طبيعيا وتباين المجتمع مجهول، فإن إحصاءة الاختبار المناسبة في هذه الحالة هي:

$$T = \frac{\bar{X} - \mu}{\sqrt{S^2/n}}$$

وتوزيعها الاحتمالي هو توزيع **t** بدرجات حرية (**V=n-1=16 - 1=15**).

وحيث أن الاختبار ذو طرف أيسر ومستوى المعنوية $\alpha = 0.10$ فمن جدول (م.**2**) نحصل على قيمة t التي على يمينها مساحة قدرها **0.10**، وهي:

$t_\alpha(v) = t_{0.10}(15) = 1.341$، والقيمة الحرجة في هذا المثال هي قيمة t التي على يسارها مساحة قدرها **0.10**، وبما أن منحنى t متماثل، إذن القيمة الحرجة تساوي **1.341−** وتكون المساحة على يسار هذه القيمة تمثل منطقة الرفض والمساحة على يمينها تمثل منطقة القبول ، وذلك كما هو واضح في الشكل **(6−8)**.



شكل (6−8)

ثم نحسب القيمة المشاهدة لإحصاءة الاختبار حيث:

$$\mathbf{t}=\frac{\overline{\mathbf{X}}-\boldsymbol{\mu}}{\sqrt{\frac{S^2}{\mathbf{n}}}}=\frac{\mathbf{205-210}}{\sqrt{\frac{\mathbf{900}}{\mathbf{16}}}}=\frac{\mathbf{(-5)(4)}}{\mathbf{30}}=-\,\mathbf{0.67}$$

وبما أن القيمة المشاهدة لإحصاءة الاختبار تقع في منطقة القبول، إذن يكون قرارنا هو قبول H_0 عند مستوى معنوية α يساوى **0.10**، ويعنى ذلك أن الوسط الحسابي لهذا المجتمع ليس أقل من **210** دينار أي أن الوسط الحسابي لدخول كل العائلات القاطنة في هذه المدينة ليس أقل من **210** دينار .

ملخص الفصل السادس

الفروض الإحصائية هي الفرع الثاني من فرعي الإحصاء الاستدلالي وهي تخمينات حول قيمة معلمة مجهولة ، وهذه التخمينات نعبِّر عنها في شكل فرضين هما :

1) فرض العدم $\mathbf{H_0}$ وهو التخمين الذي يأمل الباحث أن يرفضه .

2) الفرض البديل H_1 وهو الذي يقبل كبديل لـ H_0 عند رفض H_0 .

ويعتمد رفض H_0 أو قبوله على أساس قاعدة يضعها متخذ القرار وعلى بيانات العينة التي نحسب منها القيمة المشاهدة لإحصاءة الاختبار ، وبما أن اتخاذ القرار يعتمد على القيمة المحسوبة من العينة وقد تكون العينة لا تمثل المجتمع تمثيلا سليما مما يؤدى إلى وقوع خطأ من اثنين هما خطأ من النوع الأول ويحدث عندما تؤدى نتائج العينة إلى رفض H_0 مع أن H_0 في الواقع صحيح ونرمز لاحتماله بالرمز α ، وخطا من النوع الثاني ويحدث عندما تؤدى بيانات العينة إلى قبول H_0 مع أن H_0 في الواقع خطأ ونرمز لاحتماله بالرمز β . وإحصاءة الاختبار الخاصة باختبار الوسط الحسابي للمجتمع ، عندما يتوزع المجتمع توزيعا طبيعيا ، بتباين معلوم ، أو يكون حجم العينة كبيرا هي : $Z = \frac{\bar{X}-\mu}{\sqrt{\frac{\sigma^2}{n}}}$ وتتبع توزيع طبيعي معياري ، وعندما يتوزع المجتمع توزيعا طبيعيا ، بتباين مجهول ، فإحصاءة الاختبار هي : $T = \frac{\bar{X}-\mu}{\sqrt{\frac{S^2}{n}}}$ وتتبع توزيع t بدرجات حرية $(n-1)$.

تمارين (6)

1. عرف فرض العدم والفرض البديل.
2. تكلم عن الأخطاء التي من الممكن أن يقع فيها الباحث عند اتخاذ القرار .
3. عرف إحصاءة الاختبـــــــــــار.
4. سحبت عينة عشوائية حجمها **9** من مجتمع يتبع التوزيع الطبيعي تباينه **36**، وذلك لاختبار الفرضية التالية: $H_0 : \mu = 25$

$H_1 : \mu \neq 25$

ما هو القرار السليم الذي يجب اتخاذه عند مستوى المعنوية $\alpha = 0.10$ ، إذا علمت أن الوسط الحسابي للعينة هو **24**؟
5. إذا كان وزن إنتاج شجيرات الطماطم في إحدى المزارع له وسط حسابي$\mu = 12$ كيلو جرام، وتباين **4**، فإذا استخدم نوع جديد من السماد واختيرت عينة عشوائية تحتوى على **36** شجيرة، فوجد أن الوسط الحسابي لإنتاجها **13.5** كيلو جرام، فهل النوع الجديد من السماد أدى إلى زيادة الوسط الحسابي لإنتاج الشجيرات؟ أختبر ذلك باستخدام مستوى معنوية $\alpha = 0.05$.
6. إذا علمت أن الوقت الذي ينتظره الزبون في أحد المصارف يتبع توزيعا طبيعيا بوسط حسابي يساوى **25** دقيقة، وانحراف معياري يساوى **8** دقائق، واتبع المصرف نظاما للعمل جديدا لتقليل وقت الانتظار، وبعد اتباع هذا النظام اخترنا عينة عشوائية تشمل **100** شخص، فكان الوسط الحسابي للوقت الذي ينتظره هؤلاء الأشخاص يساوى **22** دقيقة، فأختبر فاعليه نظام العمل الجديد وذلك باستخدام مستوى معنوية $\alpha = 0.05$.

7. عينة عشوائية مسحوبة من مجتمع يتوزع توزيعا طبيعيا، فإذا علمت أن حجم العينة **25**، ووسطها الحسابي **118** وتباينها **4**، فأختبر ما إذا كان الوسط الحسابي للمجتمع المسحوبة منه هذه العينة أقل من **120**، وذلك باستخدام مستوى معنوية $\alpha = 0.01$

8. إذا علمت أن قيمة المبيعات اليومية لمحل تجاري تتبع توزيعا طبيعيا بوسط حسابي يساوى **600** دينار في اليوم، ثم أتبع هذا المحل سياسة إعلانية جديدة، وبعد إتباعها يعتقد أن الوسط الحسابي لقيمة مبيعاته اليومية قد زاد، ولاختبار ذلك اختيرت عينة عشوائية تشمل قيمة مبيعات **9** أيام، وكان الوسط الحسابي لقيمة مبيعات هذه العينة يساوى **650** دينارا، بانحراف معياري يساوى **30** دينارا، فهل كانت للسياسة الإعلانية الجديدة فاعلية؟ أختبر ذلك باستخدام مستوى معنوية $\alpha = 0.05$.

9. يدعي مدير مصنع لإنتاج نوع معين من المسامير، بأن الوسط الحسابي لطول هذا النوع من المسامير الذي ينتجه مصنعه يتبع توزيعا طبيعيا بوسط حسابي يساوي **2** سم، ولاختبار دقة إنتاج هذا المصنع اختيرت عينة عشوائية تشمل **4** مسامير وكانت أطوالها كما يلي:

1.99 ، 1.97 ، 1.98 ، 2.02

وباستخدام بيانات هذه العينة أختبر ادعاء المدير وذلك عند مستوى معنوية α يساوى **0.02** .

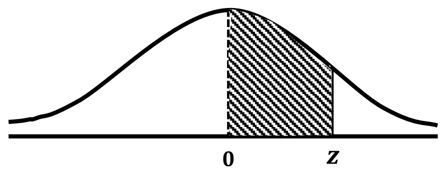
ملحق الجداول الإحصائية

جدول (م.1) المساحات تحت المنحنى الطبيعي المعياري.

جدول (م.2) القيم الحرجة لتوزيع t.

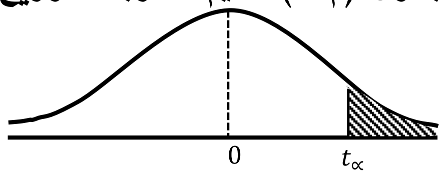
جدول (م.1)

المساحات المحصورة تحت المنحنى الطبيعي المعياري



Z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.9
0.0	.0000	.0040	.0080	.0120	.0160	.0199	.0239	.0279	.0319	.0359
0.1	.0398	.0438	.0478	.0517	.0557	.0596	.0636	.0675	.0714	.0753
0.2	.0793	.0832	.0871	.0910	.0948	.0987	.1026	.1064	.1103	.1141
0.3	.1179	.1217	.1255	.1293	.1331	.1368	.1406	.1443	.1480	.1517
0.4	.1554	.1591	.1628	.1664	.1700	.1736	.1772	.1808	.1844	.1879
0.5	.1915	.1950	.1985	.2019	.2054	.2088	.2123	.2157	.2190	.2224
0.6	.2257	.2291	.2324	.2357	.2389	.2422	.2454	.2486	.2517	.2549
0.7	.2580	.2611	.2642	.2673	.2704	.2734	.2764	.2794	.2823	.2852
0.8	.2881	.2910	.2939	.2967	.2995	.3023	.3051	.3078	.3106	.3133
0.9	.3159	.3186	.3212	.3238	.3246	.3289	.3315	.3340	.3365	.3389
1.0	.3413	.3438	.3461	.3485	.3508	.3531	.3554	.3577	.3599	.3621
1.1	.3643	.3665	.3686	.3708	.3729	.3749	.3770	.3790	.3810	.3830
1.2	.3849	.3869	.3888	.3907	.3925	.3944	.3960	.3980	.3997	.4015
1.3	.4032	.4049	.4066	.4082	.4099	.4115	.4131	.4147	.4162	.4177
1.4	.4192	.4207	.4222	.4236	.4251	.4265	.4279	.4292	.4306	.4319
1.5	.4332	.4345	.4357	.4370	.4382	.4394	.4406	.4418	.4429	.4441
1.6	.4452	.4463	.4474	.4484	.4495	.4505	.4515	.4525	.4535	.4545
1.7	.4554	.4564	.4573	.4582	.4591	.4599	.4608	.4616	.4625	.4633
1.8	.4641	.4649	.4656	.4664	.4671	.4678	.4686	.4693	.4699	.4706
1.9	.4713	.4719	.4726	.4732	.4738	.4744	.4750	.4756	.4761	.4767
2.0	.4772	.4778	.4783	.4788	.4793	.4798	.4803	.4808	.4812	.4817
2.1	.4821	.4826	.4830	.4834	.4838	.4842	.4846	.4850	.4854	.4857
2.2	.4861	.4864	.4868	.4871	.4875	.4878	.4881	.4884	.4887	.4890
2.3	.4893	.4896	.4898	.4901	.4904	.4906	.4909	.4911	.4913	.4916
2.4	.4918	.4920	.4922	.4925	.4927	.4929	.4931	.4932	.4934	.4936
2.5	.4938	.4940	.4941	.4943	.4945	.4946	.4948	.4949	.4951	.4952
2.6	.4953	.4955	.4956	.4957	.4959	.4960	.4961	.4962	.4963	.4964
2.7	.4965	.4966	.4967	.4968	.4969	.4970	.4971	.4972	.4973	.4974
2.8	.4974	.4975	.4976	.4977	.4977	.4978	.4979	.4979	.4980	.4981
2.9	.4981	.4982	.4982	.4983	.4984	.4984	.4985	.4985	.4986	.4986
3.0	.4987	.4987	.4987	.4988	.4988	.4989	.4989	.4989	.7990	.4990

جدول (م.2) القيم الحرجة لتوزيع t



Critical Values of the Distribution

ν	α				
	0.10	0.05	0.025	0.01	0.005
1	3.078	6.314	12.706	31.621	63.657
2	1.886	2.920	4.303	6.965	9.925
3	1.636	2.353	3.182	4.541	5.841
4	1.533	2.132	2.776	3.747	4.604
5	1.476	2.015	2.571	3.365	4.032
6	1.440	1.943	2.447	3.143	3.707
7	1.415	1.895	2.365	2.998	3.499
8	1.397	1.860	2.306	2.896	3.355
9	1.383	1.833	2.262	2.821	3.250
10	1.372	1.812	2.228	2.764	3.169
11	1.363	1.796	2.201	2.718	3.106
12	1.356	1.782	2.179	2.681	3.055
13	1.350	1.771	2.160	2.650	3.012
14	1.345	1.761	2.145	2.624	2.977
15	1.341	1.753	2.131	2.602	2.947
16	1.337	1.746	2.120	2.583	2.921
17	1.333	1.740	2.110	2.567	2.898
18	1.330	1.734	2.101	2.552	2.878
19	1.328	1.729	2.093	2.539	2.861
20	1.325	1.725	2.086	2.528	2.845
21	1.323	1.721	2.080	2.518	2.831
22	1.321	1.717	2.074	2.508	2.819
23	1.319	1.714	2.069	2.500	2.807
24	1.318	1.711	2.064	2.492	2.797
25	1.316	1.708	2.060	2.485	2.787
26	1.315	1.706	2.056	2.479	2.779
27	1.314	1.703	2.052	2.473	2.771
28	1.313	1.701	2.048	2.467	2.763
29	1.311	1.699	2.045	2.462	2.756
inf	1.282	1.645	1.960	2.326	2.576

قائمة أهم المراجع

1. أ. نجاة رشيد الكيخيا ، أساسيات الاحتمالات والتوزيعات الاحتمالية ، منشورات مركز بحوث العلوم الاقتصادية 2005 ف .

2. David F. Groebner & Patrick W. Shannon , Bussiness Statics (A Decision – Making Approach) . Fourth Edition , Macmillan Publishing Company . New York , 1993 .
3. Neil A . Weiss . Elementary Statistics . Second Edition , Addison – Wesley Publishing Company , 1994 .
4. Prem S. Mann , Statistics For Business And Economics , John Wiley & Sons , Inc ., New York . 1995 .
5. Robert L. Winkler & William L. Hays . Statistics (Probability , Inference , and Decision) Second Edition , Holt , Rinehart and Winston , 1975 .
6. Ronald E. Walpole , Introduction To Statistics , 2nd Edition , Macmillan Publishing Company . New York , 1974 .
7. Warren Chase & Fred Bown . General Statistics , Third Edition , John Wiley & Sons , Inc , New York , 1997 .
8. Wayne W. Daniel , Introductory Statistics With Applications , Houghton Mifflin Company . 1977 .

فهرس

قائمة المحتويات