الجمهوريــة الجزائـريــة الديمـقراطيــة الشـعبيـة
République Algérienne Démocratique et Populaire

وزارة التعليم العالي والبحث العلمي
Ministère de l'Enseignement Supérieur
et de la Recherche Scientfiique

المدرسة العليا لأساتذة التعليم التكنولوجي - سكيكدة
Ecole Normale Supérieure d'enseignement technologique, Skikda

Département de Mathématiques

قسم الرياضيات والإعلام الآلي

مطبوعة موجهة لطلبة السنة الثالثة رياضيات أستاذ التعليم الثانوي وأستاذ التعليم المتوسط بعنوان:

محاضرات في مقياس الاحتمالات والإحصاء 1

من إعداد الأستاذة :
صغير فاطمة الزهرة

السنة الجامعية: 2024-2025

الفهرس

مقدمة

الباب الأول: الإحصاء الوصفي

الباب الثاني: مدخل إلى نظرية الاحتمالات

مقدمة

تتضمن هذه المطبوعة دروس وأمثلة وفق البرنامج الوزاري والموحد للمدارس العليا للأساتذة لمقياس الاحتمالات والإحصاء 1 (Probabilités et Statistiques 1) موجهة لطلبة أقسام السنة الثالثة رياضيات أستاذ تعليم ثانوي وأستاذ تعليم متوسط، وإلى كل المستويات باختلاف تخصصاتها، وقد حرصنا في تقديم هذه المطبوعة على الإيجاز والسهولة وتبسيط المفاهيم المدعمة بأمثلة تطبيقية بسيطة بعيداً عن البراهين المعقدة والمسترسلة، وذلك حتى يتسنى للطلبة استيعاب محاور المقرر الدراسي.

وقد قُسمت هذه المطبوعة إلى بابين وكل باب مقسم بدوره إلى فصول كما يلي:

الباب الأول يتمثل في الإحصاء الوصفي الذي ينقسم إلى أربعة فصول أساسية نستهلها بالتعريف بعلم الإحصاء ثم مقاييس النزعة المركزية، يليها مقاييس التشتت، وأخيرا نتطرق إلى آخر فصل ألا وهو الارتباط والانحدار.

الباب الثاني يتمثل في مدخل إلى نظرية الاحتمالات الذي ينقسم بدوره إلى أربعة فصول أساسية نستهلها بفصل التحليل التوفيقي وأنواع السحب الذي يظم أهم التعاريف والمبادئ الأساسية في التحليل التوفيقي، نمر بعدها إلى فصل التجربة العشوائية والفضاء الاحتمالي ثم نتطرق إلى فصل المتغيرات العشوائية حيث نتعرف من خلال هذا الفصل على المتغيرات العشوائية وحيدة البعد وخصائصها وبعض قوانينها وتوزيعاتها الشهيرة وأخيرا نتطرق إلى فصل الثنائية العشوائية الذي نهتم فيه بدراسة سلوك متغيرين عشوائيين معا في آن واحد.

الباب الأول

الإحصاء الوصفي

الفصل الأول

التعريف بعلم الإحصاء

يعتبر علم الإحصاء من العلوم الضرورية والهامة لأية عملية بحث علمي أو تجربة عملية أو نظرية، أو دراسة تطبيقية تهدف إلى الوصول إلى نتائج تعتمد الموضوعية وتتسم بالمصداقية، وتكمن أهميته في شتى الميادين كونه وسيلة يستخدمها معظم الناس في حياتهم وأعمالهم اليومية، خاصة متخذي القرار في المؤسسات والإدارات لأن أي قرار اقتصادي أو إداري لا بد أن يبنى على بيانات ومعلومات دقيقة و واقعية من أجل الوصول إلى نتائج موضوعية ودقيقة عن الظاهرة أو المشكلة المراد دراستها وتحليلها.

الأهداف التعليمية للفصل:

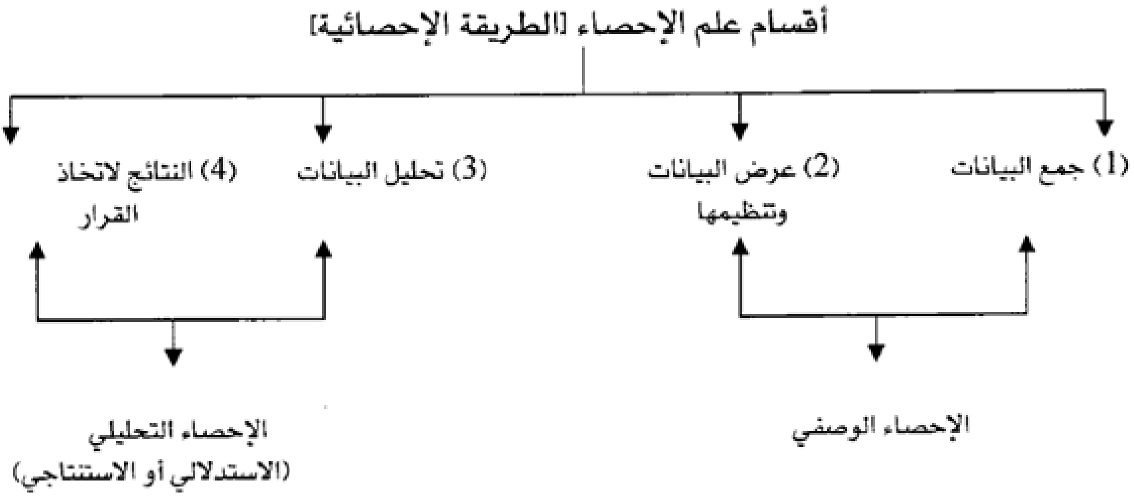
يتعرف الطالب في هذا الفصل على:

- ✓ أقسام علم الإحصاء
- ✓ المجتمع الإحصائي والعينة الإحصائية
- ✓ طرق جمع البيانات
- ✓ الجداول التكرارية البسيطة والمزدوجة
- ✓ بعض التمثيلات البيانية الهامة

1.1. علم الإحصاء

1.1.1. تعريف

الإحصاء هو العلم الذي يبحث في الأساليب والطرق العلمية المناسبة لجمع البيانات، تبويبها، تنظيمها، تحليلها وتفسيرها بهدف الوصول إلى النتائج اللازمة لزيادة المعرفة أو اتخاذ القرارات المناسبة وتعميمها.



يمكن تقسيم هذا المجال من المعرفة (علم الإحصاء) إلى نوعين رئيسيين هما:

- **الإحصاء الوصفي (Descriptive Statistics)**

يختص هذا النوع بأساليب جمع البيانات وتبويبها وعرضها، حيث تلخص البيانات ويتم اختزالها إلى معلومة أو أكثر في شكل جداول تكرارية ورسومات بيانية مع حساب بعض المقاييس الإحصائية (المتوسطات، النسب،...).

- **الإحصاء الاستقرائي (Inferential Statistics)**

يختص هذا النوع من الإحصاء باستنتاج واتخاذ القرارات المناسبة للظاهرة قيد الدراسة مع حساب درجة الثقة المصاحبة لتلك القرارات والاستنتاجات، حيث يعمم حكم الجزء على الكل.

2.1. مصطلحات إحصائية

- **المجتمع الإحصائي**

هو مجموعة العناصر أو الأفراد التي ينصب عليها الاهتمام في دراسة معينة أو مجموعة المشاهدات والقياسات الخاصة بمجموعة من الوحدات الإحصائية والتي تخص ظاهرة من الظواهر القابلة للقياس (مجتمع من الطلبة، مجتمع من الأسر،...).

وغالبا ما يكون المجتمع الإحصائي كبيرا وبالتالي دراسة جميع مفرداته قد يكون أمر غير متيسر وعليه نلجأ لدراسة جزء من مفرداته يطلق عليه تسمية عينة إحصائية.

❖ **العينة الإحصائية**

هي مجموعة جزئية من مفردات المجتمع الإحصائي، وقد جرت العادة على اختيار مفرداتها بحيث يكون لكل مفردة من مفردات المجتمع نفس الفرصة بأن تكون ضمن مفردات العينة، وذلك حتى تمثل المجتمع أحسن تمثيل، ويختلف حجم العينة حسب أهمية الدراسة وحسب الإمكانيات المادية والبشرية المتاحة للقيام بهذه الدراسة.

❖ **الوحدة الإحصائية**

هي العنصر أو الجزء الذي تجرى عليه الدراسة الإحصائية أو المعاينة، ويشترط في الوحدة أن تكون خاضعة لتعريف دقيق وواضح.

❖ **المتغيرات**

هي الصفات أو السمات التي يتصف بها أفراد عينة ما، حيث تتغير هذه الصفات من عنصر لآخر كأطوال الأطفال تحت سن معين في مدينة ما، سعر مادة معينة في سوق ما وغيرها.

❖ **المعلمة**

هو المقياس أو الثابت الذي يصف بعض خصائص المجتمع، ونحصل عليه من خلال تحليل البيانات لهذا المجتمع (متوسط دخل الفرد في بلد معين، ...)

3.1. المتغيرة الإحصائية

المتغيرة الإحصائية هي تلك الصفة أو الكمية القابلة للتغير من فرد لآخر أو من مشاهدة لأخرى والتي تسمح بتفريق هؤلاء عن أولئك وتصنفهم ويطلق على القيمة التي تعطى لها اسم قيمة المتغير الإحصائي، ويمكن تصنيف المتغيرة الإحصائية إلى قسمين:

1.3.1. المتغيرة النوعية

هي تلك المتغيرات أو الظواهر التي لا يمكن قياسها عدديا بل قياس تكرارها فقط، أو هي عبارة عن صفات أو أنواع ليست عددية وتنقسم بدورها إلى:

- **بيانات نوعية قابلة للترتيب:** يمكن ترتيبها حسب رتبة معينة تصاعديا أو تنازليا كمستويات النمو الاقتصادي، المستوى التعليمي، الرتب العسكرية، تقديرات النجاح وغيرها.
- **بيانات نوعية غير قابلة للترتيب:** كالجنسية، أنواع الأمراض، الحالة العائلية وغيرها.

2.3.1. المتغيرة الكمية

هي تلك المتغيرات التي يمكن التعبير عنها عدديا بأرقام حقيقية وقياسها رقميا وهي أكثر المتغيرات انتشارا واستعمالا لأن لغة الإحصاء هي لغة الأرقام مثال على ذلك: الإنتاج، الاستهلاك، الوزن، الطول وغيرها وتنقسم بدورها إلى قسمين:

- ✓ **المتغيرة الكمية المنفصلة أو المتقطعة أو النقطية**

هي تلك المتغيرات التي تأخذ قيما صحيحة لا يمكن تجزئتها مثل: عدد الأطفال في الأسرة، عدد الطلاب في مراحل التعليم المختلفة، عدد قطع الغيار المنتجة ... إلخ.

- ✓ **المتغيرة الكمية المتصلة أو المستمرة**

هي تلك المتغيرات التي تأخذ كل القيم الممكنة لمجال الدراسة، ونظرا للعدد غير المنتهي لهذه القيم نقسم مجال الدراسة إلى مجالات جزئية تسمى فئات.

4.1. مصادر البيانات الإحصائية

مصادر البيانات الإحصائية هي المنابع التي يأخذ منها الإحصائي البيانات موضع الدراسة، حيث يعتمد الباحثون على مصدرين أساسيين للحصول على المعلومات الإحصائية الخاصة بظاهرة معينة.

1.4.1. المصادر المباشرة

يجمع الإحصائي معلوماته إما بالاتصال والاحتكاك المباشر بوحدات المجتمع الإحصائي أو بالاعتماد على الوثائق التي تكون فيها المعلومات لازالت خامة، وفي هذه الحالة يمثل الاستبيان أو المقابلة أحد المصادر الهامة للحصول على المعلومات من المصادر الأولية.

- **الاستبانة (الاستبيان):** هي وسيلة لجمع البيانات المتعلقة بموضوع البحث عن طريق إعداد استمارة يتم تعبئتها من قبل عينة ممثلة من الأفراد.
- **المقابلة:** تعد المقابلة استبانة شفوية يقوم من خلالها الباحث (المقابل) بجمع معلومات بطريقة شفوية مباشرة من المفحوص (المستجيب) حيث أن الفرق بين المقابلة والاستبانة يكمن أن المفحوص هو الذي يكتب الإجابة على الأسئلة بينما يكتب الباحث بنفسه إجابات المفحوص في المقابلة.

2.4.1. المصادر غير المباشرة

يتحصل الباحث على المعلومات الإحصائية من الدراسات والتحقيقات السابقة، حيث تكون هذه البيانات مبوبة ومصنفة من طرف باحثين سابقين أو هيئات رسمية أو غير رسمية وتم نشرها في نشرات خاصة أو دوريات أو تكون محفوظة في الأرشيف التقليدي أو الآلي.

5.1. طرق جمع البيانات الإحصائية

عند القيام بدراسة إحصائية لظاهرة معينة يتطلب الأمر جمع بيانات ومعلومات عن وحدات المجتمع قيد الدراسة، ويتم جمع البيانات الإحصائية من الميدان باستخدام إحدى الطريقتين:

1.5.1. طريقة الحصر الشامل

يتم من خلالها تجميع البيانات من جميع المفردات التي تكون المجتمع الإحصائي كما هو الحال في التعداد العام للسكان، وتعتبر هذه الطريقة من أفضل طرق جمع البيانات وذلك لأنها تعطي بيانات كاملة حول المشكلة التي تهم الباحث، وتوفر هذه الطريقة معلومات دقيقة إذا توفرت شروط البحث العلمي، غير أنه من عيوبها ارتفاع التكاليف لكون عملية الحصر الشامل تتطلب وسائل مادية وبشرية ضخمة وتحتاج إلى وقت أطول وكلما كان حجم المجتمع كبيرا كلما كان احتمال الخطأ كبيرا، وتستخدم

هذه الطريقة إذا كان المطلوب هو الحصول على بيانات على مستوى عال من الدقة كما هو الحال في التعدادات العامة المتعلقة بالسكان، الزراعة أو الاقتصاد، ...

2.5.1. طريقة المعاينة

تستخدم هذه الطريقة إذا كان من الصعب إجراء الدراسة على كافة أفراد المجتمع أو يمكن الاكتفاء بمعلومات عن جزء من المجتمع بدلا من المجتمع ككل، ويتم اختيار جزء (عينة) من عناصر المجتمع قيد الدراسة بأسلوب علمي سليم، وبتحليل بيانات العينة إحصائيا يمكن تعميم نتائجها على المجتمع ككل، مع ملاحظة أن نتائج العينة المختارة تكون قريبة من حقائق المجتمع كلما زاد حجم العينة وكلما تم إتباع الأسلوب العلمي السليم في اختيارها.

ومن بين الأسباب التي تدعوا لإتباع طريقة المعاينة نذكر ما يلي:

- ✓ توفير المال والجهد والوقت اللازم لإجراء البحث.
- ✓ صعوبة إجراء الحصر الشامل بسبب طبيعة المجتمع، فقد يكون المجتمع غير محدود أو كبير جدا أو يتألف من وحدات ثمينة أو خطيرة أو ما شابه ذلك.

إن الهدف من المعاينة هو الوصول إلى استنتاج عن المجتمع الذي اختيرت منه العينة، وإن الخطوة الأولى في أخذ العينات هي تحديد حجم العينة، ثم البحث عن إطار المعاينة الذي سنسحب منه العينة، وبعدها نتبع إحدى إجراءات أو طرق المعاينة.

6.1. أنواع العينات

هناك أنواع عديدة من العينات ومن خلال إجراءات سحبها يمكن للباحث أن يختار ما يتناسب منها مع ظروف الدراسة التي يقوم بها، وعلى العموم يمكن تقسيم العينات إلى نوعين رئيسيين هما:

1.6.1. العينات الاحتمالية

هي العينات التي تسحب من المجتمع الإحصائي بحيث يكون لكل عنصر من عناصر المجتمع نفس الفرصة لأن يكون ضمن عناصر العينة، لذا تمتاز هذه العينة بكونها ممثلة للمجتمع الإحصائي الذي سحبت منه، وتكون قابلة لأساليب التحليل الإحصائي، ويمكن تعميم نتائجها بثقة على المجتمع الإحصائي الممثلة له، ونذكر منها:

1) العينة العشوائية البسيطة

هي العينة التي تسحب من المجتمع بحيث يكون لكل فرد من أفراد المجتمع فرصة متساوية لأن يكون ضمن العينة المختارة، ويكون هذا النوع من العينات مفيد ومؤثر في حالة وجود تجانس وصفات مشتركة بين جميع أفراد المجتمع الأصلي المعني بالدراسة.

2) العينة العشوائية الطبقية

يخص هذا النوع المجتمعات غير المتجانسة أي المجتمعات المكونة من عدة فئات اجتماعية وهذا بناءا على عدة اعتبارات منها: مستوى الدخل، مستوى الإنفاق، المستوى التعليمي وغيرها...

3) العينة العشوائية المنتظمة

تستخدم العينة العشوائية المنتظمة لاختيار عينة من مجتمع عدد عناصره محدود أو معلوم، ففي هذه الحالة نحدد المجموعة التي ستعتمد لاختيار أفراد العينة، ونبدأ باختيار رقم عشوائي من المجموعة الأولى، ثم نضيف للرقم الذي تم اختياره حجم المجموعة التي تم اعتمادها.

وبعبارة أخرى يتم اختيار وحداتها بحيث تكون المسافة أو المدة بين كل وحدة وأخرى ثابتة لجميع وحدات العينة.

4) العينة العشوائية العنقودية

وهي عينة تؤخذ للضرورة أكثر منها للاختيار، حيث يقسم المجتمع الإحصائي إلى مجموعات جزئية واضحة يسمى كل منها عنقودا، ومن كل ناحية نسحب عينة عشوائية بسيطة من بين تلك العناقيد.

2.6.1. العينات غير الاحتمالية

هناك طرق عديدة من طرق اختيار العينات غير الاحتمالية تختلف باختلاف اتجاهات الباحثين القائمين بالدراسة، غير أنه سنكتفي بعرض نوعين هما:

1) العينة العرضية

يعتمد في اختيارها على المصادفة المحضة، وتمتاز هذه الطريقة بتوفير الوقت والتكاليف، كما يمكن من خلالها الحصول على معلومات موثوقة إذا كان المجتمع المستهدف بالدراسة على جانب كبير

من التجانس، أما إذا كانت عناصر المجتمع غير متجانسة فذلك قد يؤدي إلى التحيز في اختيار العناصر المشكلة للعينة.

2) العينة الحصصية

يتم تحديد حصة مقررة لكل مجموعة أو طبقة من طبقات المجتمع المدروس، ثم تستخدم طريقة المصادفة في اختيار مفردات العينة، وتمتاز هذه الطريقة في كونها تخفض التحيز المحتمل وقوعه في العينات غير الاحتمالية، كما أنها تكون عمليا مفيدة في حالة عدم توفر أثر سحب العينات لطبقات المجتمع، لكنها تحمل مخاطرة التحيز عندما يتحقق التوازن بين حصة الطبقات من عناصر العينة.

7.1. أنواع البحوث الإحصائية

تقسم البحوث الإحصائية إلى ثلاث أقسام:

1) البحوث الوصفية

يتم من خلالها جمع المعلومات عن ظاهرة معينة لا لخدمة هدف محدد بذاته وإنما يقتصر توفير بيانات من الممكن أن تستخدم في المستقبل فيما بعد مثل: تعداد السكان، التعداد الزراعي، والصناعي....إلخ.

2) البحوث الإحصائية التحليلية

يتم بواسطتها جمع المعلومات التي تخدم هدف معين أو تساعد في تفسير مشكلة معينة لاحظها الباحث.

3) البحوث الإحصائية التجريبية

يستخدم هذا النوع من البحوث في ميادين مختلفة كالطب، الزراعة والنواحي الاقتصادية والاجتماعية. وتختلف الطرق والخطوات المستخدمة في البحوث من مجال لآخر، إلا أنها في جميع المجالات تشترك في الكثير من الجوانب، فعند القيام بأي دراسة إحصائية يجب مراعاة الخطوات التالية:

- **تحديد الظاهرة المدروسة:** حيث يتم تحديد الإطار العام للظاهرة المدروسة والذي يشمل الهدف من الدراسة، المجتمع الإحصائي، والمكان والوقت المناسب لجمع البيانات عن الظاهرة المدروسة، والصفات المطلوب معرفتها ووحدات القياس المستخدمة.

- ❖ **جمع البيانات الإحصائية:** يعتبر جمع البيانات الإحصائية من أساسيات العمل الإحصائي، حيث أن توفر البيانات الإحصائية الدقيقة والسليمة عن الظاهرة المدروسة يعطي نتائج موثوق فيها، ويساعد على اتخاذ القرارات السليمة بناء على النتائج المتوصل إليها، وحسب الهدف من الدراسة فإن الباحث يقوم بتحديد مصادر البيانات والطرق والأساليب التي يعتمد عليها للحصول على هذه البيانات.
- ❖ **تبويب وعرض البيانات:** بعد عملية جمع البيانات الإحصائية التي يمكن أن تكون أعدادها كبيرة، فذلك لا يعطينا فكرة واضحة عن نتائج جمع البيانات، لذا يلجأ الإحصائي إلى تصنيف وتبويب هذه البيانات عن طريق وضعها في مجموعات متجانسة تشترك في صفة واحدة أو عدة صفات، ويقدمها في شكل جداول أو أشكال مناسبة حتى تصبح هذه المعلومات عملية يستخدمها الباحث في حد ذاته أو تستخدم من طرف باحثين آخرين، بعد أن يتم تقديمها في نشريات خاصة أو دوريات عامة.
- ❖ **تحليل البيانات واستقراء النتائج:** تعتبر عملية تحليل البيانات مرحلة مهمة في أي بحث إحصائي وذلك لغرض الإجابة على إشكالية البحث، لذا فإن الباحث يسعى إلى التحليل الإحصائي لجوانب الظاهرة المدروسة عن طريق استخدام الأدوات الإحصائية المناسبة لتحليل البيانات من أجل الحصول على نتائج الدراسة واستقراء واستخلاص مدلولها واتخاذ القرارات على أساس النتائج المتوصل إليها.

8.1. طرق عرض البيانات الإحصائية

الخطوة التالية بعد جمع البيانات هو تبويبها وعرضها بصورة يمكن الاستفادة منها في وصف الظاهرة محل الدراسة، وهناك طريقتين لعرض البيانات وهما:

1.8.1. عرض البيانات جدوليا

يمكن عرض البيانات في صورة جداول تكرارية، ويختلف شكل الجدول طبقا لنوع البيانات، وحسب عدد المتغيرات.

1) **الجدول التكراري البسيط:** هو جدول يتكون من سطرين، أحدهما به مجموعات المتغير، والثاني به التكرارات لكل مجموعة، ولكن قبل عرضها في الجدول التكراري يجب ترتيبها تصاعديا أو تنازليا أو إدراجها ضمن فئات...إلخ.

وتختلف طريقة العرض الجدولي حسب نوع البيانات (المتغير) ونمير الحالات التالية:

❖ **الجدول التكراري البسيط للبيانات النوعية**

لتكوين جدول توزيع تكراري للبيانات النوعية نحتاج إلى إعداد جدول مكون من سطرين، يخصص السطر الأول للمتغيرات (صفات أو أنواع) بعد ترتيبها إن كانت قابلة للترتيب والسطر الثاني يخصص للتكرارات.

مثال: البيانات التالية تمثل فصائل الدم لخمسون شخص:

B، A، B، O، A، B، O، A، O، A، A، O، B، A، O، AB، A

B، A، B، O، A، B، A، A، O، A، B، O، B، A، AB، A

B، A، B، O، A، B، O، A، O، A، B، O، B، A، O، AB، A

نقوم بعرض هذه البيانات في جدول تكراري بسيط كما يلي:

فصيلة الدم	A	B	AB	O	المجموع
التكرار	20	14	3	13	50

❖ **الجدول التكراري البسيط للبيانات الكمية المنفصلة (النقطية)**

يحتوي الجدول التكراري في هذه الحالة على سطرين، الأول يمثل قيم الظاهرة X_i بعد ترتيبيا والثاني يمثل التكرارات n_i المقابلة لكل قيمة من قيم الظاهرة المدروسة، ويكون الجدول التكراري البسيط كما يلي:

X_i	X_1	X_2	...	...	...	X_k	المجموع
n_i	n_1	n_2	...	...	...	n_k	$\sum_{i=1}^{k} n_i$

مثال: فيما يلي بيانات درجات 10 طلاب في الامتحان النهائي لمقياس الإحصاء الوصفي (العلامات على 20 نقطة).

12، 10، 8، 13، 12، 8، 14، 8، 10، 12

يكون الجدول التكراري البسيط لهذه البيانات بعد ترتيبها تصاعديا كما يلي:

العلامة	8	10	12	13	14	المجموع
التكرار	3	2	3	1	1	10

❖ الجدول التكراري البسيطة للبيانات الكمية المستمرة (المتصلة)

هي أكثر الجداول استخداماً ويمكن أن تأخذ مفرداتها أرقاما صحيحة أو كسرية. عند دراسة بيانات كمية مستمرة فإن مجال الدراسة يضم عدد كبير من القيم ولتعذر وضع كل هذه القيم نقسم هذا المجال إلى مجالات جزئية تسمى فئات (Classes)، حيث يتم تحديد عدد الفئات حسب حجم العينة المدروسة من جهة وحسب توزيع الوحدات الإحصائية من جهة أخرى، ولتشكيل الجدول التكراري الخاص بالبيانات الكمية المتصلة نتبع الخطوات التالية:

- **تحديد المدى العام:** يرمز لو بالرمز E وهو المجال الذي تنتشر فيه البيانات إذ يمثل الفرق بين أكبر مشاهدة وأصغر مشاهدة، وعليه فإن:

$$E = X_i(\max) - X_i(\min)$$

- **تحديد عدد الفئات:** غالبا ما يتراوح عدد الفئات في التوزيع التكراري بين 5 و 20 فئة ويعتمد هذا على عدد المفردات في العينة، يتم تحديد عدد الفئات اللازمة لتشكيل الجدول التكراري باستخدام بعض المعادلات الرياضية، أهمها على الإطلاق هي:
  - قاعدة ستيرج (La règle de Sturge) وتعطى بالصيغة التالية:

$$K = 1 + 3.322\log(N)$$

  - قاعدة يول (La règle de Yule) وتعطى بالصيغة التالية:

$$K = 2.5 \times \sqrt[4]{N}$$

- **تحديد طول الفئة:** يتم تحديد طول الفئة بالعلاقة التالية:

$$L = \frac{E}{K}$$

- **تحديد حدود الفئات:** في هذه المرحلة يتم تحديد بداية ونهاية كل فئة، على أن تكون بداية الفئة الأولى أصغر من أو تساوي أصغر قيمة في البيانات ونهاية الفئة الأخيرة أكبر من أكبر قيمة في البيانات.

مثال: فيما يلي بيانات تمثل توزيع درجات الأجور الشهرية (10^3 دج) التي حصل عليها 42 عامل في مؤسسة ما:

26	23	55	40	56	52	66
14	64	51	44	38	42	49
38	42	16	67	63	30	26
45	67	42	35	53	15	14
40	50	45	50	57	56	60
49	52	35	39	53	46	24

لكي يتم تبويب هذه البيانات في شكل جدول تكراري، يتم إتباع ما يلي:

- ✓ **حساب المدى (E):**

$$E = X_i(\max) - X_i(\min) = 67 - 14 = 53$$

- ✓ **تحديد عدد الفئات (K):** باستخدام قاعدة ستيرج، نجد أن

$$K = 1 + 3.322\log(N) = 1 + 3.322\log(42) = 6.39 \approx 6$$

- ✓ **حساب طول الفئة (L):**

$$L = \frac{E}{K} = \frac{53}{6} = 8.83 \approx 9$$

- ✓ **تحديد الفئات:**

- الحد الأدنى للفئة الأولى هو الحد الأدنى للسلسلة = 14
- الحد الأعلى للفئة الأولى = الحد الأدنى + طول الفئة = 14+ L =14 + 9= 23.
- الحد الأدنى للفئة الثانية هو الحد الأعلى للفئة الأولى = 23.
- الحد الأعلى للفئة الثانية = الحد الأدنى + طول الفئة = 160+ L =23 + 9 = 32.

وبنفس الطريقة يتم تكوين حدود الفئات الأخرى، وبالتالي يكون الجدول التكراري كما يلي:

فئات الأجر	$[14;23[$	$[23;32[$	$[32;41[$	$[41;50[$	$[50;59[$	$[59;68[$	المجموع
التكرارات	4	5	7	7	11	8	42

ملاحظة: في الجداول التكرارية للبيانات الكمية المتصلة (ذات فئات) يجب مراعاة أن يكون لكل قيمة فئة واحدة وواحدة فقط والتأكد من أن مجموع التكرارات يساوي عدد القيم.

كما يمكن إثراء الجدول التكراري البسيط من خلال إضافة أسطر أخرى خاصة بكل من:

- ❖ **التوزيع التكراري النسبي** (f_i)**:** بعد إعداد جدول التوزيع التكراري يكون من المناسب في أغلب الأحيان عرض البيانات في شكل توزيع تكراري نسبي للتعبير عن الأهمية النسبية لتكرار كل قيمة بالنسبة لإجمالي التكرارات، ويحسب التكرار النسبي بالصيغة التالية:

$$f_i = \frac{n_i}{\sum_{i=1}^{k} n_i} = \frac{n_i}{N}$$

- ❖ **التوزيع التكراري المئوي** $f_i(\%)$ **:** يمكن الحصول عليه بضرب التكرار النسبي في 100. ويلاحظ أن مجموع التكرارات المئوية يساوي 100. وتعطى علاقة التكرار النسبي المئوي بالصيغة التالية:

$$f_i(\%) = f_i \times 100$$

- ❖ **التوزيع التكراري المتجمع الصاعد** $(n_i \uparrow)$**:** يعرف التكرار المتجمع الصاعد (ت.م.ص.) على أنه مجموع التكرارات التي هي أقل من أو تساوي تكرار قيمة ما، وعليه يمكن حسابه كما يلي:

X_i	X_1	X_2	X_3	...	...	X_k	المجموع
n_i	n_1	n_2	n_3	...	...	n_k	$\sum_{i=1}^{k} n_i$
ت.م.ص. $n_i \uparrow$	n_1	n_1+n_2	$n_1+n_2+n_3$	...	...	$\sum_{i=1}^{k} n_i$	

❖ **التوزيع التكراري المتجمع النازل** $(n_i \downarrow)$ **:** يعرف التكرار المتجمع النـازل (ت.م.ن.) علـى أنـه مجموع التكرارات التي هي أكبر من أو تساوي تكرار قيمة ما، وعليه يمكن حسابه كما يلي:

X_i	X_1	...	...	X_{k-2}	X_{k-1}	X_k	المجموع
n_i	n_1	...	...	n_{k-2}	n_{k-1}	n_k	$\sum_{i=1}^{k} n_i$
ت.م.ن. $n_i \downarrow$	$\sum_{i=1}^{k} n_i$	...	...	$n_k + n_{k-1} + n_{k-2}$	$n_k + n_{k-1}$	n_k	

مثـال: باسـتخدام معطيـات المثـال السـابق، نقـوم بإيجـاد التكـرارات النسـبية والنسـبية المئويـة وكـذا التكرارات المتجمعة الصاعدة والنازلة كما يلي:

فئات الأجر X_i	$[14;23[$	$[23;32[$	$[32;41[$	$[41;50[$	$[50;59[$	$[59;68[$	$\sum n_i$
n_i	4	5	7	7	11	8	42
f_i	0.095	0.12	0.17	0.17	0.26	0.19	1
$f(\%)_i$	9.5	12	17	17	26	19	100
$n_i \uparrow$	4	4+5=9	9+7=16	+7=23 16	+11=34 23	+8=42 34	
$n_i \downarrow$	42	42-4=38	38-5=33	33-7=26	26-7=19	19-11=8	
$n_i \downarrow$	+4=42 38	+5=38 33	+7=33 26	+7=26 19	8+11=19	8	

2) **جداول التوزيع التكراري المزدوجة:** يستعمل جدول التوزيـع التكراري المزدوج عند دراسة ظاهرتين أو خاصيتين معا في آن واحد، حيث توضع البيانات الإحصائية جدوليا على الشكل التالي:

✓ تخصـــص الأســـطر لبيانـــات الخاصـــية الأولـــى، ونرمـــز لقـــيم الخاصـــية الأولـــى بــــ $X_i = \{X_1, X_2, ..., X_p\}$،

- ✓ تخصص الأعمدة لبيانات الخاصية الثانية، ونرمز لبيانات الخاصية الثانية بـ $Y_j = \{Y_1, Y_2, ..., Y_q\}$ ،
- ✓ نضع $N = n_{ij}$ عدد الأفراد في المجتمع.

X_i \ Y_j	Y_1	Y_2	Y_3	...	X_q	المجموع $n_{i.}$
X_1	n_{11}	n_{12}	n_{13}	...	n_{1q}	$n_{1.}$
X_2	n_{21}	n_{22}	n_{23}	...	n_{2q}	$n_{2.}$
X_3	n_{31}	n_{32}	n_{33}	...	n_{3q}	$n_{3.}$
...	...	...	...	...	...	...
X_p	n_{p1}	n_{p2}	n_{p3}	...	n_{pq}	$n_{p.}$
المجموع $n_{.j}$	$n_{.1}$	$n_{.2}$	$n_{.3}$	...	$n_{.q}$	$n_{ij} = \sum_{i=1}^{p} n_{i.} = \sum_{j=1}^{q} n_{.j}$

نرمز لمحصلة الجمع على المعامل i أو المعامل j بنقطة.

مثال: سحبت عينة عشوائية من مجتمع ما تتكون من 100 أسرة قصد دراسة ظاهرتين معا في آن واحد هما:

- ✓ النفقات الاستهلاكية للأسرة (10^3 دج)،
- ✓ التركيبة الأسرية من حيث عدد الأطفال.

فكانت النتائج المتحصل عليها مبينة في الجدول التالي:

عدد الأطفال \ النفقات	0	1	2	3	المجموع
[20;30[	3	8	6	3	20
[30;40[	8	15	12	13	48
[40;50[	4	7	11	10	32
المجموع	15	30	29	26	100

نلاحظ من خلال الجدول أن:

- ✓ توجد 32 عائلة تتراوح نفقاتهم الاستهلاكية بين 40×10^3 دج و 50×10^3 دج.
- ✓ توجد 26 عائلة لديها 3 أطفال، حيث أن 3 عائلات منها تتراوح نفقاتهم الاستهلاكية بين 20×10^3 دج و 30×10^3 دج، و 13 عائلات منها تتراوح نفقاتهم الاستهلاكية بين 30×10^3 دج و 40×10^3 دج، بينما 10 عائلات الأخرى تتراوح نفقاتهم الاستهلاكية بين 40×10^3 دج و 50×10^3 دج.

2.8.1. عرض البيانات بيانيا

يفيد التمثيل البياني للتوزيعات التكرارية في إظهار الخصائص الرئيسية مما يساعد على متابعة التغيير في الظاهرة والطريقة التي يتغير بها توزيع ما من فترة إلى أخرى، وتستخدم أنواع مختلفة للعرض البياني حسب نوع المتغير المدروس.

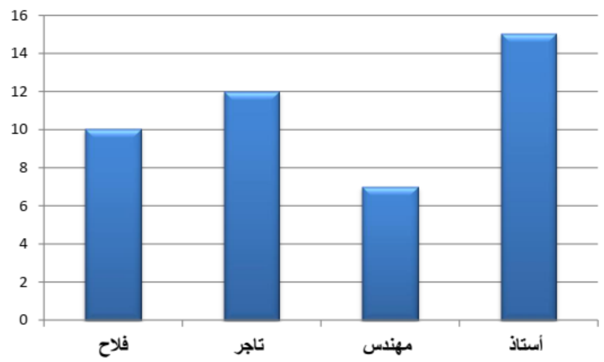
1) الأعمدة البيانية البسيطة

- **في حالة البيانات النوعية**

هي عبارة عن مستطيلات متباعدة بمسافات ثابتة ولها قواعد متساوية تتناسب أطوالها مع تكرارات الخاصية المدروسة.

مثال: نقوم بتمثيل معطيات الجدول التالي عن طريق الأعمدة البيانية البسيطة.

المهنة	فلاح	تاجر	مهندس	أستاذ
عدد الأشخاص	10	12	7	15



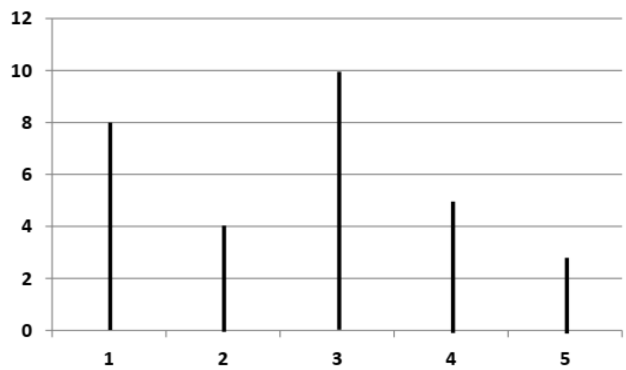
في حالة البيانات الكمية المنفصلة (النقطية)

هي عبارة عن خطوط مستقيمة نقطية طولها يعبر عن تكرارات الظاهرة المدروسة.

مثال: نقوم بتمثيل بيانات معطيات الجدول التكراري التالي.

X_i	1	2	3	4	5
n_i	8	4	10	5	3

الأعمدة البيانية لهذه البيانات تعطى كما يلي:



2) المدرج التكراري

هو التمثيل البياني للجدول التكراري البسيط الخاص بالبيانات الكمية المستمرة (المتصلة)، وهو عبارة عن مستطيلات متلاصقة، حيث تمثل التكرارات على محور رأسي، بينما تمثل حدود الفئات على محور أفقي، ويتم تمثيل كل فئة بمستطيل، ارتفاعه هو تكرار الفئة، وطول قاعدته هو طول الفئة.

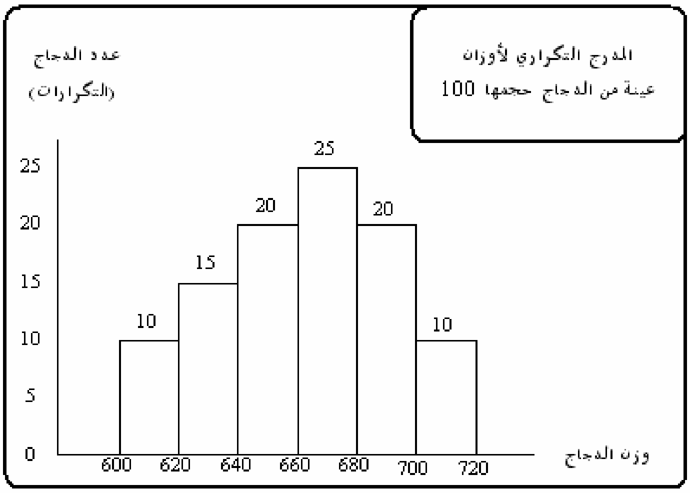
المدرج التكراري في حالة فئات متساوية الطول

عندما تكون الفئات متساوية الطول فإن قاعدة المقارنة ثابتة ومتساوية وعليه نقوم برسم المدرج التكراري مباشرة.

مثال: فيما يلي التوزيع التكراري لأوزان عينة من الدواجن بالغرام، عددها 100 اختيرت من أحد المزارع بعد 45 يوم.

الوزن X_i	[600;620[	[620;640[	[640;660[	[660;680[	[680;700[	[700;720[	المجموع
عدد الدجاج n_i	10	15	20	25	20	10	100

المدرج التكراري لهذه البيانات يعطى كما يلي:



المدرج التكراري في حالة فئات غير متساوية الطول

إذا كانت فئات التوزيع غير متساوية الطول نقوم بتعديل التكرارات لأن قاعدة المقارنة غير ثابتة، حتى يكون هناك تناسب بين طول الفئة والتكرار المقابل لها، ولغرض تعديل التكرارات نستخدم المعادلة الآتية:

$$\text{التكرار المعدل} = n_i^* = \frac{n_i}{L_i}$$

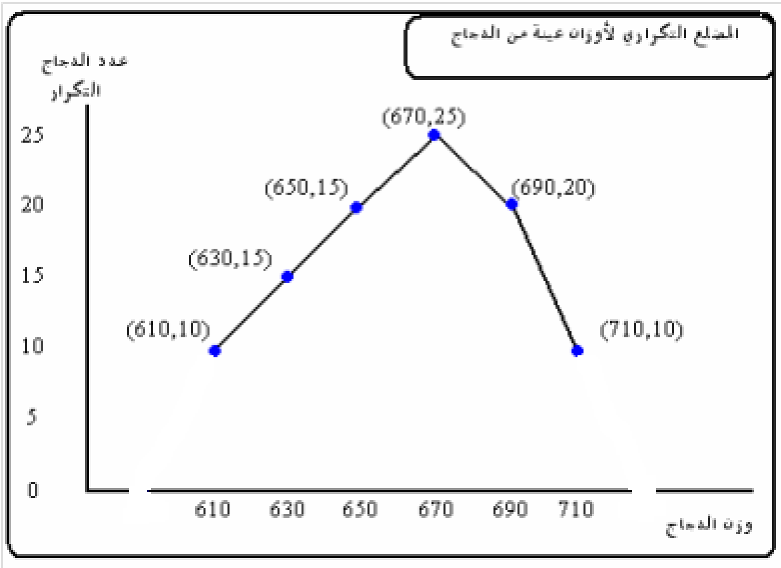
3) المضلع التكراري

لإنشاء المضلع التكراري نقوم بتعيين مجموعة النقط $M_i(C_i, n_i)$ في معلم متعامد ومتجانس، ثم نوصل بينها بخط منكسر. حيث:

- ✓ C_i هي مركز الفئة و $C_i = \frac{a+b}{2}$،
- ✓ a الحد الأدنى للفئة،
- ✓ b الحد الأعلى للفئة.

مثال: نستخدم بيانات الجدول التكراري السابق لرسم المضلع التكراري.

الوزن X_i	[600;620[	[620;640[	[640;660[	[660;680[	[680;700[	[700;720[	المجموع
مراكز الفئات C_i	$\frac{600+620}{2}=610$	$\frac{620+640}{2}=630$	$\frac{640+660}{2}=650$	$\frac{660+680}{2}=670$	$\frac{680+700}{2}=690$	$\frac{700+720}{2}=710$	
عدد الدجاج n_i	10	15	20	25	20	10	100



4) المنحنى التكراري

بإتباع نفس الخطوات السابقة في رسم المضلع التكراري، يمكن رسم المنحنى التكراري، ولكن نوصل بين النقاط بخط منحنى.

5) الدائرة النسبية

هي عبارة عن دائرة مقسمة إلى عدة أجزاء كل جزء يقابل زاوية مركزية تتناسب مع التكرارات المقابلة لكل خاصية من الخصائص المدروسة، حيث نقوم بإضافة سطر إلى جدول المعطيات يحتوي على الزوايا المركزية المقابلة لكل تكرار.

يتم حساب الزوايا المركزية باستخدام العلاقة التالية:

$$\text{الزاوية المركزية} = f_i \times 360^\circ$$

مثال: الجدول التالي يمثل التوزيع النسبي المئوي لليد العاملة في مختلف القطاعات في مرحلتين مختلفتين

القطاع / السنة	1960	1985
الفلاحة	62	25
الصناعة	6	14
البناء	4	17
الإدارة	17	28
الخدمات	11	16
مجموع النسب	%100	%100

نقوم بتمثيل بيانات الجدول في دائرة نسبية كما يلي:

الفصل الثاني

مقاييس النزعة المركزية

في الكثير من النواحي التطبيقية يكون الباحث في حاجة إلى حساب بعض المؤشرات التي يمكن الاعتماد عليها في وصف الظاهرة من حيث القيمة التي تتوسط القيم أو تنزع إليها القيم، ومن حيث التعرف على مدى تجانس القيم التي يأخذها المتغير، وأيضا ما إذا كان هناك قيم شاذة أم لا. والاعتماد على العرض البياني وحده لا يكفي، لذا يتناول هذا الفصل عرض بعض المقاييس الإحصائية التي يمكن من خلالها التعرف على خصائص الظاهرة محل البحث.

الأهداف التعليمية للفصل:

يتعرف الطالب في هذا الفصل على:

- ✓ مفهوم مقاييس النزعة المركزية
- ✓ المتوسط الحسابي
- ✓ الوسيط
- ✓ المنوال
- ✓ المتوسط الهندسي
- ✓ الربيعيات

2. مقاييس النزعة المركزية

تسمى مقاييس النزعة المركزية بمقاييس الموضع أو المتوسطات، وهى القيم التي تتركز القيم حولها، ومن هذه المقاييس، الوسط الحسابي، الوسيط، المنوال، المتوسط الهندسي، وغيرها من المقاييس الإحصائية الأخرى، وفيما يلي عرض لبعض وأهم هذه المقاييس.

1.2. المتوسط الحسابي

يعتبر المتوسط الحسابي $(\overline{X})$ أو الوسط الحسابي أكثر مقاييس النزعة المركزية استخداما، فهو مركز توازن لأي ظاهرة. ويعرف الوسط الحسابي بشكل عام على أنه مجموع القيم مقسوما على عددها، ونميز الحالات التالية:

1.1.2. حالة قيم مفردة دون تكرار

إذا كان لدينا n من القيم المفردة والغير مكررة، $x_1, x_2, ..., x_n$. نعرف الوسط الحسابي في هذه الحالة بـ:

$$\overline{X} = \frac{\sum_{i=1}^{n} x_i}{N}$$

حيث N هو عدد القيم.

2.1.2. حالة قيم مفردة بتكرار

إذا كان لدينا مجموعة من القيم لظاهرة معينة $x_1, x_2, ..., x_k$ وكانت $n_1, n_2, ..., n_k$ التكرارات المقابلة لها، ننظم هذه القيم في جدول كما يلي

القيم x_i	x_1	x_2	.	.	x_k	المجموع
التكرار n_i	n_1	n_2	.	.	n_k	$N = \sum_{i=1}^{k} n_i$

نعرف عندئذ الوسط الحسابي بـ:

$$\overline{X}=\frac{\sum_{i=1}^{k}n_i x_i}{N}=\sum_{i=1}^{k}f_i x_i$$

حيث $f_i=\frac{n_i}{N}$ هو التواتر النسبي للسلسلة الإحصائية.

3.1.2. حالة جداول تكرارية ذات الفئات

في هذه الحالة يكون جدول التوزيع التكراري على شكل فئات عددها k، بعد حساب مراكز الفئات $C_1,C_2,...,C_k$، ننظم القيم في جدول كما يلي:

الفئات	$[a_1,a_2[$	$[a_2,a_3[$	…	…	$[a_k,a_{k+1}[$	
المراكز C_i	C_1	C_2	…	…	C_k	المجموع
التكرارات n_i	n_1	n_2	…	…	n_k	$N=\sum_{i=1}^{k}n_i$

حيث $C_i=\frac{a_{k-1}+a_k}{2}$ هو مركز الفئة، نعرف عندئذ الوسط الحسابي بـ:

$$\overline{X}=\frac{\sum_{i=1}^{k}x_i C_i}{N}=\sum_{i=1}^{k}f_i C_i$$

مثال: الجدول التالي يمثل توزيع 40 طالب حسب أوزانهم.

الوزن	$[44,46[$	$[46,48[$	$[48,50[$	$[50,52[$	$[52,54[$	$[54,56[$
عدد التلاميذ	4	7	13	10	5	1

الوسط الحسابي لوزن الطلبة هو:

$$\overline{X}=\frac{\sum_{i=1}^{6}x_iC_i}{N}=\frac{45\times4+47\times7+49\times13+51\times10+53\times5+55\times1}{40}=\frac{1976}{40}=49,4$$

4.1.2. خواص المتوسط الحسابي

- يعتبر المتوسط الحسابي من أهم مقاييس النزعة المركزية وأكثرها استخداما.
- المتوسط الحسابي للمقدار الثابت يساوي الثابت نفسه.
- اذا أضيف مقدار ثابت إلى كل قيمة من القيم، فان المتوسط الحسابي للقيم بعد الإضافة يساوي المتوسط الحسابي للقيم الأصلية مضافا إليها هذا المقدار الثابت، أي أن: $\overline{X+a}=\overline{X}+a$.
- اذا ضرب مقدار ثابت في كل قيمة من القيم، فان المتوسط الحسابي للقيم الناتجة يساوي المتوسط الحسابي للقيم الأصلية مضروبا في هذا المقدار الثابت، أي أن: $\overline{X\times a}=\overline{X}\times a$.
- يتأثر المتوسط الحسابي بالقيم المتطرفة في الصغر والكبر وينحاز لصالحها ومن ثم فلا يعتمد عليه.
- مجموع انحرافات القيم عن وسطها الحسابي معدوم $\sum_{i=}^{k}\left(x_i-\overline{X}\right)=0$.

2.2. الوسيط

تعريف 1.2.2.

الوسيط (Med) هو قيمة المتغير الإحصائي الذي يفصل السلسلة الإحصائية إلى قسمين متساويين بعد ترتيب القيم تصاعديا أو تنازليا.

تعريف 2.2.2.

الوسيط هو القيمة x_i ، حيث 50% من التكرار الكلي تأخذ قيم أكبر من قيمة الوسيط و 50% المتبقية تأخذ قيمة أصغر من قيمة الوسيط.

1.2.2. حساب الوسيط في حالة القيم المنفصلة (النقطية)

بعد ترتيب السلسلة تصاعديا أو تنازليا، نميز حالتين:

1) إذا كان N فردي ($N=2p+1$)، عندئذ يكون الوسيط هو القيمة التي رتبتها $p+1$ ، ونكتب:

$$Med=x_{p+1}$$

2) إذا كان N زوجي ($N=2p$)، عندئذ يكون الوسيط هو نصف مجموع القيمتان اللتان رتبتاهما p و $p+1$ ، تعطى عبارة الوسيط في هذه الحالة بـ:

$$Med=\frac{x_p+x_{p+1}}{2}$$

مثال 1: أوجد الوسيط للقيم التالية: 50، 40، 45، 40، 55، 135، 60

نرتب السلسلة ترتيبا تصاعديا: 40، 40، 45، 50، 55، 60، 135

$N=7=2\times3+1$ ، وبالتالي يكون الوسيط هو:

$$Med=x_{3+1}=x_4=50$$

مثال 2: أوجد الوسيط للقيم التالية: 41، 95، 42، 43، 45، 34

نرتب السلسلة ترتيبا تصاعديا: 34، 41، 42، 43، 45، 95

$N=6=2\times3$ ، وبالتالي يكون الوسيط هو:

$$Med=\frac{x_3+x_{3+1}}{2}=\frac{x_3+x_4}{2}=\frac{42+43}{2}=42{,}5$$

2.2.2. حساب الوسيط في الحالة المتصلة (المستمرة)

لحساب الوسيط في حالة سلسلة إحصائية مستمرة نتبع الخطوات التالية:

1) نكون جدول التوزيع المتجمع الصاعد.

2) نحسب رتبة الوسيط وهي $\frac{N}{2}$ بغض النظر عما إذا كان N زوجي أو فردي.

3) نعين الفئة التي تكرارها المجتمع الصاعد $\frac{N}{2}$ أو أكثر وتسمى الفئة الوسطية.

4) نستعمل القانون التالي:

$$Med=a+\frac{\frac{N}{2}-n_c}{n_e}\times l$$

حيث أن:

- a الحد الأدنى للفئة الوسطية،

- $\frac{N}{2}$ رتبة الوسيط في السلسلة الإحصائية،
- n_c تكرار المتجمع الصاعد للفئة قبل الوسطية،
- n_e تكرار الفئة الوسطية،
- l طول الفئة الوسطية.

مثال: نقوم بإيجاد الوسيط للسلسلة الإحصائية التالية

الفئات	$[20,25[$	$[25,30[$	$[30,35[$	$[35,40[$	$[40,45[$	$[45,50[$	$[50,55[$	≥ 55
n_i	3	9	13	16	20	15	13	11
$n\uparrow$	3	12	25	41	61	76	89	100

- رتبة الوسيط في السلسلة الإحصائية هي $\frac{N}{2}=\frac{100}{2}=50$
- الفئة الوسطية هي $[40,45[$ وبالتالي:

$$l=5 \text{ ، } n_e=20 \text{ ، } n_c=41 \text{ ، } a=40$$

وبتطبيق القانون نجد أن:

$$Med=40+\frac{50-41}{20}\times 5=42,25$$

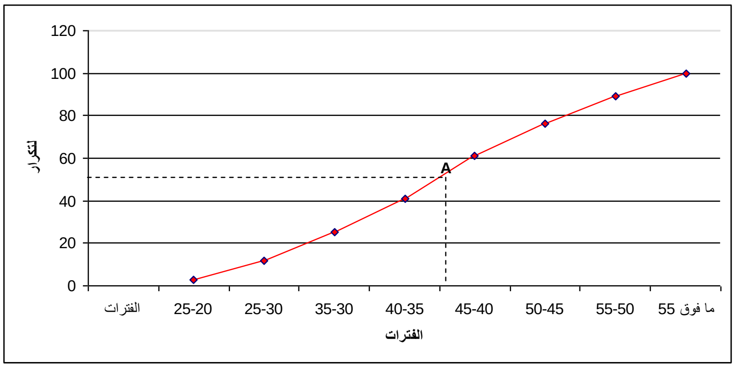
3.2.2. حساب الوسيط بيانيا

طريقة 1:

لحساب الوسيط بيانيا نتبع الخطوات التالية:

1) نكون جدولا لتوزيع التكرار المتجمع الصاعد.
2) نرسم منحنى التكرار المجتمع الصاعد.
3) نحدد رتبة الوسيط على المحور العمودي، ونرسم منها مستقيم موازيا للمحور الأفقي، فيقطع منحنى المتجمع الصاعد في نقطة فاصلتها هي الوسيط.

مثال: باستخدام المثال السابق نجد أن الوسيط بيانيا هو



من خلال الرسم نلاحظ أن فاصلة A هي تقريبا $Med = 42{,}3$.

طريقة 2:

الوسيط بيانيا هو نقطة تقاطع كل من منحنى التكرار المتجمع الصاعد والنازل، فالإسقاط العمودي لنقطة التقاطع بين المنحنيين على محور الفواصل يعطينا قيمة الوسيط.

4.2.2. خواص الوسيط

- يصف البيانات بأكثر واقعية في حالة وجود قيم شاذة أو متطرفة ولا يتأثر بها.
- يمكن حسابه من الجداول التكرارية ذات الفترات غير محدودة.
- غير قابل للعمليات الجبرية.
- لا يدخل في حسابه جميع البيانات المعطاة بل يتأثر بترتيبها فقط.
- يمكن حسابه عن طريق الرسم.
- ليس له معنى في حالة عدم ترتيب البيانات.

3.2. المنوال

المنوال (Mode) هو القيمة التي تحمل أكبر تكرار، وقد يكون للسلسلة الإحصائية أكثر من منوال واحد.

1.3.2. حساب المنوال في حالة القيم المنفصلة

مثال: السلسلة التالية تمثل نتائج 30 تلميذ في امتحان مقياس الإحصاء.

15	13	12	10	7	القيم(العلامات)
2	7	8	8	5	التكرارات

نلاحظ أنه يوجد أكثر من منوال واحد لهذه السلسلة، أي أن:

$$Mod = 10,\ 12$$

2.3.2. حساب المنوال في الحالة المستمرة (المتصلة)

نسمي الفئة التي لها أكبر تكرار بالفئة المنوالية ويقع المنوال بداخلها ولحساب المنوال نتبع الخطوات التالية:

1) نحدد الفئة المنوالية وليكن تكرارها f ،
2) نعين f_1 تكرار الفئة قبل المنوالية،
3) نعين f_2 تكرار الفئة بعد المنوالية،
4) نحسب $\Delta_1 = f - f_1$ و $\Delta_2 = f - f_2$ ،
5) نطبق القانون التالي:

$$Mod = a + \frac{\Delta_1}{\Delta_1 + \Delta_2} \times l$$

حيث a الحد الأدنى للفئة المنوالية، l هو طول الفئة المنوالية.

مثال: أوجد المنوال للسلسلة الإحصائية التالية

الفئات	[54,57[	[57,60[	[60,63[	[63,66[	[66,69[	[69,72[	[72,75[	[75,78[	[78,81[
التكرارات	3	4	10	15	15	17	12	5	3

نلاحظ أن أكبر تكرار هو $f = 17$ ، وبالتالي الفئة المنوالية هي [69,72[

وبالتالي

$$a = 69 \text{ ، } f_1 = 15 \text{ ، } f_2 = 12 \text{ ، } \Delta_1 = 2 \text{ ، } \Delta_2 = 5 \text{ ، } l = 3$$

بتطبيق القانون نجد أن

$$Mod = 69 + \frac{2}{2+5} \times 3 = 69,85$$

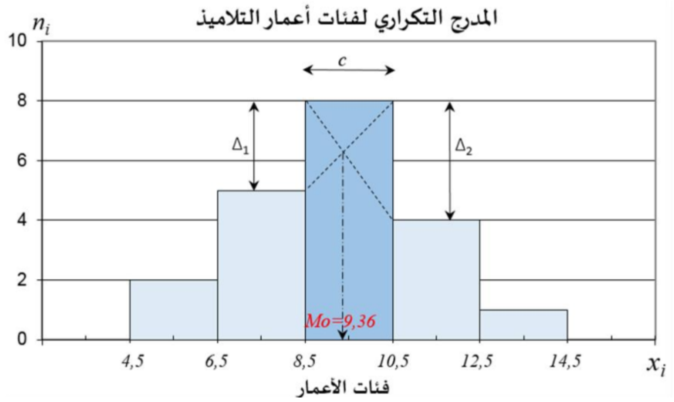
3.3.2. حساب المنوال عن طريق الرسم

يمكن حساب المنوال عن طريق الرسم وذلك باستخدام المدرج التكراري للفئة المنوالية والتي قبلها والتي بعدها ثم نصل نهاية المستطيل للفئة المنوالية من الناحية اليسرى بنهاية المستطيل للفئة التي بعدها من الجهة اليسرى، كذلك نهاية المستطيل الفئة المنوالية من الناحية اليمنى بنهاية المستطيل الذي قبله من اليمين ومن نقطة التقاطع ننزل عمود على المحور الأفقي، فنحصل على المنوال.

مثال: نقوم بإيجاد المنوال بيانيا لبيانات تلاميذ النادي الموسيقي التالية.

فئات الأعمار	[4.5,6.5[	[6.5,8.5[	[8.5,10.5[	[10.5,12.5[	[12.5,14.5[	$\sum$
n_i التكرار	2	5	8	4	1	20

بعد رسم المدرج التكراري لهذه السلسلة، يكون المنوال بيانيا كما هو موضح



4.3.2. خواص المنوال

- لا يتأثر بالقيم المتطرفة.
- يمكن حسابه من الجداول التكرارية ذات فترات غير محدودة.
- يعتبر من أحسن المتوسطات لوصف الظواهر النوعية.
- يمكن حسابه عن طريق الرسم ويمكن أن يكون لدينا أكثر من منوال واحد.
- لا يدخل في حسابه جميع البيانات.

4.2. المتوسط الهندسي

1.4.2. حالة قيم مفردة دون تكرار

إذا كان لدينا n من القيم، و يرمز لها بـ: $x_1, x_2,......, x_n$ ، يعرف المتوسط الهندسي في هذه الحالة على أنه الجذر النوني لجداء هذه القيم وهو معطى بالعلاقة التالية:

$$G=\sqrt[n]{x_1\times x_2\times\ldots\times x_n}$$

ولغرض تسهيل العملية الحسابية في حال كون البيانات كبيرة فإنه يفضل استخدام أخذ لوغاريتم البيانات لتصبح صيغة المتوسط الهندسي كما يلي:

$$G=\sqrt[n]{x_1\times x_2\times\ldots\times x_n}=\left(x_1\times x_2\times\ldots\times x_n\right)^{\frac{1}{n}}$$

بإدخال اللوغاريتم على الطرفين نتحصل على

$$\ln G=\ln\left(x_1\times x_2\times\ldots\times x_n\right)^{\frac{1}{n}}=\frac{1}{n}\ln\left(x_1\times x_2\times\ldots\times x_n\right)=\frac{1}{n}\left(\ln\left(x_1\right)+\ln\left(x_2\right)+\ldots+\ln\left(x_n\right)\right)=\frac{1}{n}\sum_{i=1}^{n}\ln\left(x_i\right)$$

وبالتالي تصبح عبارة المتوسط الهندسي كما يلي:

$$G=e^{\frac{1}{n}\sum_{i=1}^{n}\ln\left(x_i\right)}$$

2.4.2. حالة قيم مفردة بتكرار

إذا كان لدينا مجموعة من القيم لظاهرة معينة $x_1,x_2,\ldots,x_k$ وكانت $n_1,n_2,\ldots,n_k$ التكرارات المقابلة لها. يعطى المتوسط الهندسي في هذه الحالة بالعلاقة التالية:

$$G=\sqrt[N]{x_1^{\,n_1}\times x_2^{\,n_2}\times\ldots\times x_k^{\,n_k}}=\left(x_1^{\,n_1}\times x_2^{\,n_2}\times\ldots\times x_k^{\,n_k}\right)^{\frac{1}{N}}$$

حيث $N=\sum_{i=1}^{k}n_i$

بإدخال اللوغاريتم على الطرفين تصبح علاقة المتوسط الهندسي كما يلي:

$$G=e^{\frac{1}{N}\sum_{i=1}^{n}n_i\ln\left(x_i\right)}$$

3.4.2. حالة الجداول التكرارية ذات الفئات

في هذه الحالة نقوم بحساب مراكز الفئات $C_1,C_2,\ldots,C_k$، ويعطى المتوسط الهندسي في هذه الحالة بالعلاقة التالية:

$$G=\sqrt[N]{C_1^{\,n_1}\times C_2^{\,n_2}\times\ldots\times C_k^{\,n_k}}=\left(C_1^{\,n_1}\times C_2^{\,n_2}\times\ldots\times C_k^{\,n_k}\right)^{\frac{1}{N}}$$

حيث $N=\sum_{i=1}^{k} n_i$

بإدخال اللوغاريتم على الطرفين تصبح علاقة المتوسط الهندسي كما يلي:

$$G=e^{\frac{1}{N}\sum_{i=1}^{n} n_i \ln(C_i)}$$

مثال: نقوم بحساب المتوسط الهندسي للبيانات التالية (المثال السابق) لتلاميذ النادي الموسيقي.

فئات الأعمار	$[4.5,6.5[$	$[6.5,8.5[$	$[8.5,10.5[$	$[10.5,12.5[$	$[12.5,14.5[$	$\sum$
C_i	5.5	7.5	9.5	11.5	13.5	
التكرار n_i	2	5	8	4	1	20
$\ln(C_i)$	1.7047	2.0149	2.2513	2.4423	2.6027	
$n_i \times \ln(C_i)$	3.4095	10.0745	18.0103	9.7694	2.6027	43.8664

وبالتالي يكون المتوسط الهندسي هو:

$$G=e^{\frac{1}{20}\sum_{i=1}^{5} n_i \ln(C_i)}=e^{\frac{43.8664}{20}}=8.9649$$

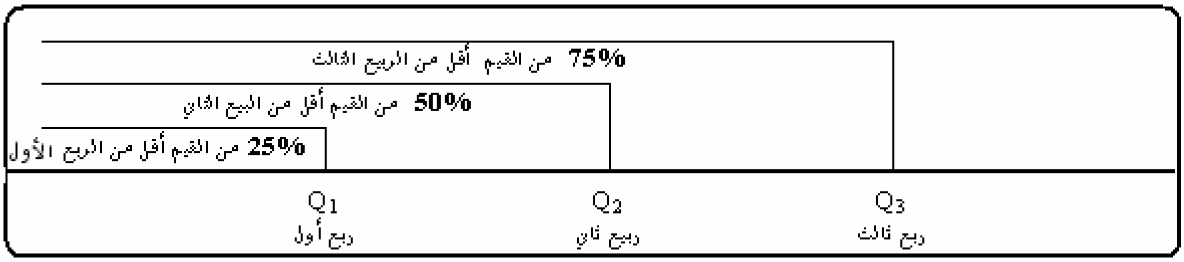
4.4.2. خواص الوسط الهندسي

- يدخل في حسابه جميع القيم ولكنه أقل تأثرا بالقيم المتطرفة من المتوسط الحسابي.
- لا يمكن حسابه من جداول التوزيع التكراري المفتوحة.
- ليس له معنى إذا كانت إحدى القيم سالبة أو تساوي الصفر.
- يستخدم بأكثر واقعية عند وصف الظواهر النسبية.
- قيمة المتوسط الهندسي لأي ظاهرة أصغر دائما من قيمة المتوسط الحسابي.

5.2. الربيعيات

الربيعيات (Les quartiles) هي قيم تأخذها المتغيرة الإحصائية بحيث تقسم التكرار الكلي إلى نسب معينة:

- ❖ **الربيعي الأول** (Q_1) لسلسلة إحصائية هو أصغر قيمة للمتغيرة الإحصائية حيث يكون 25% على الأقل من الحدود لها قيمة أصغر أو تساوي Q_1.
- ❖ **الربيعي الثاني** (Q_2) يتطابق مع الوسيط أي أن $Q_2 = Med$.
- ❖ **الربيعي الثالث** (Q_3) لسلسلة إحصائية هو أصغر قيمة للمتغيرة الإحصائية حيث يكون 75% على الأقل من الحدود لها قيمة أصغر أو يساوي Q_3.



1.5.2. حساب الربيعيات في الحالة متقطعة

لحساب الربيعيات في حالة سلسلة إحصائية متقطعة نتبع الخطوات التالية:

1) ترتيب القائمة ترتيبا تصاعديا.

2) نحدد قيمة Q_1، القيمة التي رتبتها n حيث n هو أصغر عدد طبيعي يحقق $n \geq \frac{N}{4}$.

3) نحدد قيمة Q_3، القيمة التي رتبتها n حيث n هو أصغر عدد طبيعي يحقق $n \geq \frac{3N}{4}$.

مثال: نحسب الربيعيات للسلسلة الإحصائية التالية:

القيم X_i	4	7	10	13	16
التكرار n_i	6	5	7	4	1
ت.م.ص. $n\uparrow$	6	11	18	22	23

نلاحظ أن:

$$\frac{N}{4}=\frac{23}{4}=5,75 \quad \rightarrow \quad Q_1=4$$

$$\frac{N}{2}=\frac{23}{2}=11,5 \quad \rightarrow \quad Q_2=10$$

$$\frac{3N}{4}=\frac{3\times 23}{4}=17,25 \quad \rightarrow \quad Q_3=10$$

2.5.2. حساب الربيعيات في الحالة المستمرة

لحساب الربيعيات في حالة سلسلة إحصائية مستمرة نتبع ما يلي:

1) **الربيعي الأول:** نستخدم القانون التالي:

$$Q_1=a^{'}+\frac{N/4-n_c^{'}}{n_1}\times l^{'}$$

حيث أن:

- $a^{'}$ هي الحد الأدنى للفئة التي يقع فيها الربيعي الأول،
- $\frac{N}{4}$ رتبة الربيعي الأول،
- $n_c^{'}$ تكرار المتجمع الصاعد للفئة قبل فئة الربيعي الأول،
- n_1 تكرار فئة الربيعي الأول،
- $l^{'}$ هو طول فئة الربيع الأول.

2) **الربيعي الثاني:** يتطابق مع الوسيط.

3) **الربيعي الثالث:** نستخدم بالقانون التالي:

$$Q_3=a^{''}+\frac{3N/4-n_c^{''}}{n_3}\times l^{''}$$

حيث أن:

- a'' هي الأدنى للفئة التي يقع فيها الربيع الأول،
- $\frac{3N}{4}$ رتبة الربيع الثالث،
- $n_c^{''}$ تكرار المتجمع الصاعد للفئات قبل فئة الربيع الثالث،

- n_3 تكرار فئة الربيع الثالث،
- l'' هو طول فئة الربيع الثالث.

مثال: نقوم بحساب الربيعيات للبيانات التالية لتلاميذ النادي الموسيقي.

فئات الأعمار	$[4.5,6.5[$	$[6.5,8.5[$	$[8.5,10.5[$	$[10.5,12.5[$	$[12.5,14.5[$	$\sum$
C_i	5.5	7.5	9.5	11.5	13.5	
التكرار n_i	2	5	8	4	1	20
ت.م.ص. $n\uparrow$	2	7	15	19	20	

لحساب الربيعيات في هذه الحالة نتبع ما يلي:

1) الربيعي الأول:

رتبة الربيعي الأول في السلسلة الإحصائية هي $\frac{N}{4}=\frac{20}{4}=5$ وبالتالي تكون فئة الربيعي الأول هي $[6.5,8.5[$.

باستخدام القانون نجد أن:

$$Q_1 = a' + \frac{N/4 - n_c'}{n_1} \times l' = 6.5 + \frac{5-2}{7} \times 2 = 7.7$$

2) الربيعي الثالث:

رتبة الربيعي الثالث في السلسلة الإحصائية هي $\frac{3N}{4}=\frac{60}{4}=15$ وبالتالي تكون فئة الربيعي الثالث هي $[8.5,10.5[$.

باستخدام القانون نجد أن:

$$Q_3 = a'' + \frac{3N/4 - n_c''}{n_3} \times l'' = 8.5 + \frac{15-7}{8} \times 2 = 10.5$$

الفصل الثالث

مقاييس التشتت

إن مقاييس النزعة المركزية وحدها لا تكفي لإعطاء صورة كاملة عن علاقة البيانات ببعضها البعض، فقد نجد أن لسلسلتين مختلفتين نفس المتوسط الحسابي، بينما مدى البيانات للسلسلتين مختلف، هذا الفرق في البيانات مقابل تساوي المتوسط الحسابي يجعل من الضرورة استخدام مقاييس أخرى تكمل المقاييس الأولى، وتسمى هذه المقاييس بمقاييس التشتت.

الأهداف التعليمية للفصل:

يتعرف الطالب في هذا الفصل على:

- ✓ مفهوم مقاييس التشتت
- ✓ المدى
- ✓ الانحراف الربيعي
- ✓ الانحراف المتوسط
- ✓ التباين والانحراف المعياري

3. مقاييس التشتت

تشتت بيانات ظاهرة ما يقصد به درجة أو مقدار التفاوت أو الاختلاف بين مفردات هذه الظاهرة، وتعتبر بيانات الظاهرة متجانسة عندما تكون قيمها قريبة من بعضها البعض ونقول في هذه الحالة أنها غير مشتتة، أما إذا كانت بيانات الظاهرة متباعدة وغير متجانسة فنقول أن مفردات (قيم) الظاهرة مشتتة وغير مركزة.

1.3. المدى

المدى (Etendue) هو أبسط مقاييس التشتت مفهوما وتطبيقا وحسابا، يعطي المدى فكرة عن انتشار قيم السلسلة بشرط عدم وجود قيم متطرفة. ونميز الحالتين التاليتين:

1.1.3. في حالة السلاسل الإحصائية المنفصلة (المتقطعة)

المدى في حالة السلاسل الإحصائية المنفصلة هو عبارة عن الفرق بين أكبر قيمة وأصغر قيمة للسلسة الإحصائية.

ونكتب:

$$E = X_{Max} - X_{Min}$$

2.1.3. في حالة السلاسل الإحصائية ذات الفئات

يكون لدينا في هذه الحالة أكثر من صيغة نذكر منها:

المدى= مركز الفئة الأخيرة – مركز الفئة الأولى

المدى= الحد الأعلى للفئة الأخيرة – الحد الأدنى للفئة الأولى

3.1.3. خواص المدى

- بسيط الحساب وسهل الفهم ويعتمد في حسابه على قيمتين فقط.
- شديد التأثر بالقيم المتطرفة.
- لا يمكن حسابه من جداول التوزيع التكراري المفتوحة.

2.3. الانحراف الربيعي (المدى الربيعي)

للتخلص من بعض عيوب المدى والتي من أهمها تأثره بالقيم المتطرفة وعدم إمكانية حسابه في حالة جداول التوزيع التكراري المفتوحة، وجد مقياس آخر وهو المدى الربيعي حيث نقوم في هذا المقياس بإهمال الربع الأول والربع الأخير من البيانات المرتبة تصاعديا، وفي هذه الحالة تكون أكبر قيمة في البيانات هي الربيع الثالث وأصغر قيمة في البيانات هي الربيع الأول، والفرق بينهما يعطي ما يسمى بالمدى الربيعي.

الانحراف الربيعي (E_Q) وهو عبارة عن الفرق بين الربيعي الثالث والربيعي الأول، وهذا المقياس يبين لنا المجال الذي ينتشر فيه نصف البيانات. ونكتب:

$$E_Q = Q_3 - Q_1$$

مثال: نحسب الانحراف الربيعي للسلسلة الإحصائية التالية

الفئات	$[8,12[$	$[12,16[$	$[16,20[$	$[20,24[$	$[24,28[$	$[28,32[$	$\sum$
التكرارات	3	13	18	38	22	6	100
$n_i \uparrow$	3	16	34	72	94	100	

نقوم أولا بحساب الربيعي الأول والربيعي الثالث كما يلي:

- ✓ رتبة الربيعي الأول $\frac{N}{4} = 25$ وبالتالي فئة الربيعي الأول هي $[16,20[$.

بتطبيق قانون الربيعي الأول نجد أن:

$$Q_1 = 16 + \frac{25-16}{18} \times 4 = 18$$

- ✓ رتبة الربيعي الثالث $\frac{3N}{4} = 75$ وبالتالي فئة الربيعي الثالث هي $[24,28[$.

بتطبيق قانون الربيعي الثالث نجد أن:

$$Q_3 = 24 + \frac{75-72}{22} \times 4 = 24{,}545$$

وبالتالي يكون الانحراف الربيعي هو:

$$E_Q = Q_3 - Q_1 = 24{,}545 - 18 = 6.545$$

1.2.3. خواص الانحراف الربيعي

- لا يتأثر بالقيم المتطرفة ويمكن حسابه بيانيا،
- يمكن حسابه في حالة جداول التوزيع التكراري المفتوحة،
- يتحدد بعدد البيانات وليس بقيمها،
- عبارة عن مجال يحتوي على 50% من البيانات.

3.3. الانحراف المتوسط

الانحراف المتوسط $(E_{\overline{X}})$ هو أحد مقاييس التشتت، ويعبر عنه بمتوسط الانحرافات المطلقة للقيم عن وسطها الحسابي. ونميز حالتين:

1.3.3. حالة قيم مفردة بتكرار

إذا كان لدينا مجموعة من القيم لظاهرة معينة $x_1, x_2, ..., x_k$ وكانت $n_1, n_2, ..., n_k$ التكرارات المقابلة لها. يعطى الانحراف المتوسط في هذه الحالة بالعلاقة التالية:

$$E_{\overline{X}} = \sum_{i=1}^{k} \frac{n_i \left| x_i - \overline{X} \right|}{N} = \sum_{i=1}^{k} \frac{n_i}{N} \left| x_i - \overline{X} \right| = \sum_{i=1}^{k} f_i \left| x_i - \overline{X} \right|$$

مثال: لتكن السلسلة الإحصائية التالية

القيم x_i	3	5	7	9	11	13
التواتر f_i	0.08	0.15	0.28	0.35	0.1	0.04

- نحسب أولا الوسط الحسابي للسلسلة كما يلي:

$$\overline{X} = \sum_{i=1}^{6} f_i x_i = 3 \times 0.08 + 5 \times 0.15 + 7 \times 0.28 + 9 \times 0.35 + 11 \times 0.1 + 13 \times 0.04 = 7.72$$

- نضيف سطرا آخر للجدول لحساب $\left|x_i - \overline{X}\right|$ كما يلي:

القيم x_i	3	5	7	9	11	13
التواتر f_i	0.08	0.15	0.28	0.35	0.1	0.04
$\left\|x_i - \overline{X}\right\|$	4.72	2.72	0.72	1.28	3.28	5.28
$f_i\left\|x_i - \overline{X}\right\|$	0.37	0.40	0.20	0.44	0.32	0.21

- وأخيرا نحسب المجموع:

$$E_{\overline{X}} = \sum_{i=1}^{k} f_i\left|x_i - \overline{X}\right| = 0.37 + 0.40 + 0.20 + 0.44 + 0.32 + 0.21 = 1.94$$

2.3.3. حالة السلاسل الإحصائية المستمرة

في هذه الحالة نقوم بحساب مراكز الفئات $C_1, C_2, ..., C_k$ ، ويعطى الانحراف المتوسط في هذه الحالة بالعلاقة التالية:

$$E_{\overline{X}} = \sum_{i=1}^{k} \frac{n_i\left|C_i - \overline{X}\right|}{N} = \sum_{i=1}^{k} \frac{n_i}{N}\left|C_i - \overline{X}\right| = \sum_{i=1}^{k} f_i\left|C_i - \overline{X}\right|$$

2.3.3. خواص الانحراف المتوسط

- يتأثر بالقيم المتطرفة ويعتمد في حسابه على جميع القيم،
- لا يمكن حسابه في حالة جداول التوزيع التكراري المفتوحة،
- يمكن حسابه عن طريق الانحرافات عن الوسيط مع الملاحظة أن الانحراف المتوسط عن الوسيط أقل من الانحراف المتوسط عن المتوسط الحسابي.

4.3. التباين والانحراف المعياري

1.4.3. التباين

التباين $Var(X)$ هو عبارة عن المتوسط الحسابي لمربعات الفروق بين قيم المتغير الإحصائي ومتوسطها الحسابي.

1) **في حالة سلسلة إحصائية متقطعة:** لتكن $\{x_1, x_2, x_3, ...x_n\}$ سلسلة إحصائية، يعطى التباين في هذه الحالة بالعبارة التالية

$$Var(X) = \frac{1}{N}\sum_{i=1}^{k} n_i\left(x_i - \overline{X}\right)^2$$

ويمكن كتابة التباين في الحالة النقطية (المنفصلة) بعدة صيغ كما يلي:

$$Var(X) = \frac{1}{N}\sum_{i=1}^{k} n_i\left(x_i - \overline{X}\right)^2 = \sum_{i=1}^{k} f_i\left(x_i - \overline{X}\right)^2 = \frac{1}{N}\sum_{i=1}^{k} n_i x_i^2 - \overline{X}^2$$

2) **في حالة السلاسل الإحصائية المستمرة:** في هذه الحالة نقوم بحساب مراكز الفئات $C_1, C_2, ..., C_k$، ويعطى التباين في هذه الحالة بالعلاقة التالية:

$$Var(X) = \frac{1}{N}\sum_{i=1}^{k} n_i\left(C_i - \overline{X}\right)^2$$

ويمكن كتابة التباين في حالة السلاسل المستمرة (المتصلة) بعدة صيغ كما يلي:

$$Var(X) = \frac{1}{N}\sum_{i=1}^{k} n_i\left(C_i - \overline{X}\right)^2 = \sum_{i=1}^{k} f_i\left(C_i - \overline{X}\right)^2 = \frac{1}{N}\sum_{i=1}^{k} n_i C_i^2 - \overline{X}^2$$

2.4.3. الانحراف المعياري

يعرف الانحراف المعياري على أنه الجذر التربيعي للتباين، ويعتبر من أهم المقاييس الإحصائية للتشتت وأكثرها استخداما في النظريات والقوانين الإحصائية، ويرمز له بـ σ_X، ويعطى بالعبارة التالية:

$$\sigma_X = \sqrt{Var(X)}$$

مثال: نحسب الانحراف المعياري لبيانات أعمار التلاميذ كما يلي

فئات الأعمار	[4.5,6.5[	[6.5,8.5[	[8.5,10.5[	[10.5,12.5[	[12.5,14.5[	$\sum$
C_i	5.5	7.5	9.5	11.5	13.5	
n_i التكرار	2	5	8	4	1	20
$n_i \times C_i$	11	37.5	76	46	13.5	184
C_i^2	30,25	56,25	90,25	132,25	182,25	
$n_i \times C_i^2$	60,5	281,25	722	529	182,25	1775

المتوسط الحسابي لهذه السلسلة هو

$$\overline{X} = \frac{\sum_{i=1}^{5} x_i C_i}{20} = \frac{184}{20} = 9.2 \quad \rightarrow \quad \overline{X}^2 = 84.64$$

تباين هذه السلسلة الإحصائية هو:

$$Var(X) = \frac{1}{N}\sum_{i=1}^{k} n_i C_i^2 - \overline{X}^2 = \frac{1775}{20} - 84.64 = 4.11$$

وبالتالي يكون الانحراف المعياري هو:

$$\sigma_X = \sqrt{Var(X)} = \sqrt{4.11} = 2.0273$$

3.4.3. خواص الانحراف المعياري

- لا يمكن حسابه من جداول التوزيع التكراري المفتوحة يتأثر بكل القيم وخاصة المتطرفة منها،
- قابل للعمليات الجبرية،
- كثير الاستعمال في القوانين والنظريات الإحصائية،
- إذا كانت لدينا القيم $x_1, x_2, ... , x_n$ ، وإذا تم:
  - ✓ إضافة أو طرح قيمة ثابتة a لكل قيمة x_i فان الانحراف المعياري لا يتغير،

✓ ضرب أو قسمة كل قيمة في مقدار ثابت (موجب) فان الانحراف المعياري يتأثر بالمقدار نفسه. ونكتب:

$$\sigma(x \mp a) = \sigma(x) \qquad \sigma\left(\frac{X}{c}\right) = \frac{1}{c}\sigma(x) \qquad c > 0$$

$$\sigma(\alpha X) = \alpha\sigma(x) \quad \alpha > 0 \qquad \sigma\left(\frac{X \mp a}{c}\right) = \frac{1}{c}\sigma(x)$$

➢ دور الانحراف المعياري هو معرفة كيفية انتشار القيم حول الوسط الحسابي.

الفصل الرابع

الارتباط والانحدار

في الفصول السابقة تم عرض بعض المقاييس الوصفية كمقاييس النزعة المركزية ومقاييس التشتت والتي يمكن من خلالها وصف شكل توزيع البيانات التي تم جمعها من متغير واحد. في هذا الفصل ننتقل من التعامل مع متغير واحد إلى التعامل مع متغيرين، وذلك باستخدام بعض طرق التحليل الإحصائي مثل تحليل الارتباط والانحدار الخطي البسيط، فإذا كان اهتمام الباحث هو دراسة العلاقة بين متغيرين استخدم لذلك أسلوب تحليل الارتباط أما إذا كان اهتمامه دراسة أثر أحد المتغيرين على الآخر استخدم لذلك أسلوب تحليل الانحدار، ومن الأمثلة على ذلك:

- ✓ القامة والوزن عند مجموعة من الأشخاص.
- ✓ القامة والعمر عند مجموعة من الأطفال.
- ✓ المردود الفلاحي و كمية الأسمدة.

الأهداف التعليمية للفصل:

يتعرف الطالب في هذا الفصل على:

- ✓ مفهوم الارتباط والانحدار
- ✓ التوزيعات الهامشية
- ✓ الخصائص الإحصائية الهامشية
- ✓ التغاير
- ✓ معامل الارتباط الخطي البسيط
- ✓ مستقيم الانحدار

4. الارتباط والانحدار

إن معرفة طبيعة واتجاه العلاقة بين المتغيرات يدعى الارتباط، والذي من خلاله يمكننا التنبؤ بظاهرة أو موقف ما من خلال ما يعرف بعملية الانحدار. ولاشك أن الارتباط والانحدار وجهان يكملان بعضهما البعض.

1.4. السلاسل الإحصائية المزدوجة

عرفنا في الفصل الأول السلاسل الإحصائية المزدوجة وذكرنا أنها تستعمل عند دراسة ظاهرتين أو خاصيتين معا في آن واحد، حيث توضع البيانات الإحصائية في جدول على الشكل التالي:

- ✓ تخصص الأسطر لبيانات الخاصية الأولى، ونرمز لقيم الخاصية الأولى بـ $X_i = \{x_1, x_2, ..., x_p\}$ ،
- ✓ تخصص الأعمدة لبيانات الخاصية الثانية، ونرمز لبيانات الخاصية الثانية بـ $Y_j = \{y_1, y_2, ..., y_q\}$ ،
- ✓ نضع N عدد الأفراد في المجتمع.

$X_i \backslash Y_j$	y_1	y_2	y_3	...	y_q
x_1	n_{11}	n_{12}	n_{13}	...	n_{1q}
x_2	n_{21}	n_{22}	n_{23}	...	n_{2q}
x_3	n_{31}	n_{32}	n_{33}	...	n_{3q}
...	...	...	...	...	...
x_p	n_{p1}	n_{p2}	n_{p3}	...	n_{pq}

1.1.4. التوزيعات الهامشية

تعريف 1.1.4.

في كلتا الحالتين (المستمرة و المتقطعة)، تحسب المجاميع الخاصة بسطر من السطور أو عمود من الأعمدة، فنحصل على التكرارات الهامشية $n_{.j}$, $n_{i.}$ المعرفين كما يلي:

$$n_{i.}=\sum_{j=1}^{q}n_{ij} \quad و \quad n_{.j}=\sum_{i=1}^{p}n_{ij}$$

$X_i \backslash Y_j$	y_1	y_2	y_3	...	y_q	المجموع $n_{i.}$
x_1	n_{11}	n_{12}	n_{13}	...	n_{1q}	$n_{1.}$
x_2	n_{21}	n_{22}	n_{23}	...	n_{2q}	$n_{2.}$
x_3	n_{31}	n_{32}	n_{33}	...	n_{3q}	$n_{3.}$
...	...	...	...	...	...	...
x_p	n_{p1}	n_{p2}	n_{p3}	...	n_{pq}	$n_{p.}$
المجموع $n_{.j}$	$n_{.1}$	$n_{.2}$	$n_{.3}$	...	$n_{.q}$	$N=\sum_{i=1}^{p}n_{i.}=\sum_{j=1}^{q}n_{.j}$

نرمز لمحصلة الجمع على المعامل i أو المعامل j بنقطة.

ملاحظات:

- مجموع التكرارات الهامشية يساوي التكرار الكلي N. ونكتب

$$N=\sum_{i=1}^{p}\sum_{j=1}^{q}n_{ij}=\sum_{i=1}^{p}n_{i.}=\sum_{j=1}^{q}n_{.j}$$

- مجموعة الثنائيات $(x_i,n_{i.})$ تسمى التوزيع الهامشي للمتغير X.
- مجموعة الثنائيات $(y_j,n_{.j})$ تسمى التوزيع الهامشي للمتغير Y.

وبالتالي يكون:

- جدول التوزيع الهامشي لـ X

x_i	x_1	x_2	...	x_p
$n_{i.}$	$n_{1.}$	$n_{2.}$	...	$n_{p.}$

- جدول التوزيع الهامشي لـ Y

y_j	y_1	y_2	…	y_q
$n_{.j}$	$n_{.1}$	$n_{.2}$	…	$n_{.q}$

2.1.4. الخصائص الإحصائية للتوزيع التكراري المزدوج

1) المتوسط الحسابي الهامشي

❖ **المتوسط الحسابي الهامشي للمتغير X**

هو المتوسط الحسابي لقيم X وما يقابلها من تكرارت هامشية سواء مطلقة أو نسبية، ويعطى بالعبارة التالية:

- في حالة سلسلة إحصائية متقطعة

$$\overline{X} = \frac{\sum_{i=1}^{p} n_{i.} x_i}{\sum_{i=1}^{p} n_{i.}} = \sum_{i=1}^{p} f_{i.} x_i$$

- في حالة سلسلة إحصائية مستمرة، نحسب مراكز الفئات

$$\overline{X} = \frac{\sum_{i=1}^{p} n_{i.} C_i}{\sum_{i=1}^{p} n_{i.}} = \sum_{i=1}^{p} f_{i.} C_i$$

❖ **المتوسط الحسابي الهامشي للمتغير Y**

هو المتوسط الحسابي لقيم Y وما يقابلها من تكرارت هامشية سواء مطلقة أو نسبية، ويعطى بالعبارة التالية:

- في حالة سلسلة إحصائية متقطعة

$$\bar{Y}=\frac{\sum_{j=1}^{q} n_{;j} y_i}{\sum_{j=1}^{q} n_{.j}}=\sum_{j=1}^{q} f_{.j} y_j$$

- في حالة سلسلة إحصائية مستمرة، نحسب مراكز الفئات

$$\bar{Y}=\frac{\sum_{j=1}^{q} n_{;j} C_i}{\sum_{j=1}^{q} n_{.j}}=\sum_{j=1}^{q} f_{.j} C_j$$

2) التباين الهامشي

❖ التباين الهامشي للمتغير X

هو التباين لقيم X وما يقابلها من تكرارت هامشية سواء مطلقة أو نسبية، ويعطى بالعبارة التالية:

- في حالة سلسلة إحصائية متقطعة

$$Var(X)=\frac{1}{N}\sum_{i=1}^{p} n_{i.}\left(x_i-\bar{X}\right)^2=\sum_{i=1}^{p} f_{i.}\left(x_i-\bar{X}\right)^2=\frac{1}{N}\sum_{i=1}^{p} n_{i.} x_i^2-\bar{X}^2$$

- في حالة سلسلة إحصائية مستمرة، نحسب مراكز الفئات

$$Var(X)=\frac{1}{N}\sum_{i=1}^{p} n_{i.}\left(C_i-\bar{X}\right)^2=\sum_{i=1}^{p} f_{i.}\left(C_i-\bar{X}\right)^2=\frac{1}{N}\sum_{i=1}^{p} n_{i.} C_i^2-\bar{X}^2$$

❖ التباين الهامشي للمتغير Y

هو التباين لقيم Y وما يقابلها من تكرارت هامشية سواء مطلقة أو نسبية، ويعطى بالعبارة التالية:

- في حالة سلسلة إحصائية متقطعة

$$Var(Y)=\frac{1}{N}\sum_{j=1}^{q} n_{.j}\left(y_j-\bar{Y}\right)^2=\sum_{j=1}^{q} f_{.j}\left(y_j-\bar{Y}\right)^2=\frac{1}{N}\sum_{j=1}^{q} n_{i.} y_j{}^2-\bar{Y}^2$$

- في حالة سلسلة إحصائية مستمرة، نحسب مراكز الفئات

$$Var(Y)=\frac{1}{N}\sum_{j=1}^{q}n_{.j}\left(C_j-\overline{Y}\right)^2=\sum_{j=1}^{q}f_{.j}\left(C_j-\overline{Y}\right)^2=\frac{1}{N}\sum_{j=1}^{q}n_{i.}C_j^{\ 2}-\overline{Y}^2$$

3) التباين المشترك (التغاير)

التباين المشترك للثنائية (X,Y) يعرف كما يلي:

- في حالة سلسلة إحصائية متقطعة

$$Cov(X,Y)=\frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{q}n_{ij}\left(x_i-\overline{X}\right)\left(y_j-\overline{Y}\right)$$

- في حالة سلسلة إحصائية مستمرة، نحسب مراكز الفئات

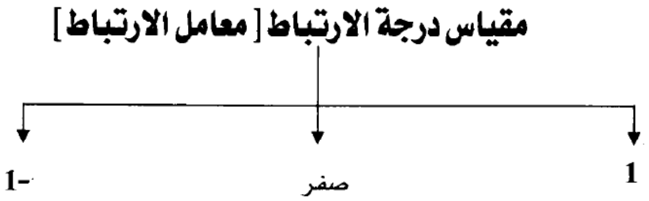
$$Cov(X,Y)=\frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{q}n_{ij}\left(C_i-\overline{X}\right)\left(C_j-\overline{Y}\right)$$

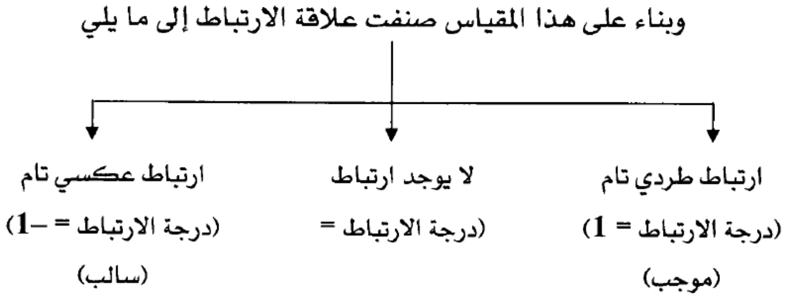
خواص التغاير:

1. $Cov(X,Y)=\frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{q}n_{ij}x_iy_j-\overline{X}\,\overline{Y}$
2. $Cov(X,X)=Var(X)$
3. إذا كان X و Y متغيرين عشوائيين مستقلين، فإن $Cov(X,Y)=0$.

2.4. معامل الارتباط

الغرض من تحليل الارتباط هو تحديد نوع و قوة العلاقة بين ظاهرتين. وأبسط الارتباطات هو الارتباط الخطي البسيط الذي يبحث في قوة العلاقة بين ظاهرتين من خلال حساب معامل الارتباط الذي نرمز له بالرمز ρ والذي نجد أنه يركز على ما يلي:





ملاحظة هامة: تزداد قوة الارتباط كلما اقتربنا من الأطراف وتقل كلما ابتعدنا عن الأطراف.

1.2.4. معامل الارتباط في حالة القيم المفردة

تعريف 1.2.4.

يعرف معامل الارتباط بـالعلاقة التالية:

$$\rho=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}=\frac{\frac{1}{N}\sum_1^p x_iy_i-\overline{X}.\overline{Y}}{\sqrt{\frac{1}{N}\sum_1^p x_i^2-\left(\frac{1}{N}\sum_1^p x_i\right)^2}\cdot\sqrt{\frac{1}{N}\sum_1^p y_i^2-\left(\frac{1}{N}\sum_1^p y_i\right)^2}}$$

حيث σ_X و σ_Y هما الانحراف المعياري لـ X و Y على الترتيب (في الحالة النقطية).

2.2.4. معامل الارتباط في حالة الجداول المزدوجة

تعريف 2.2.4.

يعرف معامل الارتباط بـالعلاقة التالية:

$$\rho = \frac{Cov(X,Y)}{\sigma_X \sigma_Y} = \frac{\frac{1}{N}\sum_{i=1}^{p}\sum_{j=1}^{q} n_{ij}(x_i - \overline{X})(y_j - \overline{Y})}{\sqrt{\frac{1}{N}\sum_{i=1}^{p} n_{i.} \cdot (x_i - \overline{X})^2}\sqrt{\frac{1}{N}\sum_{j=1}^{q} n_{.j} \cdot (y_j - \overline{Y})^2}}$$

مثال: نحسب معامل الارتباط السلسلة الإحصائية التالية:

$Y \backslash X$	$[20,30[$	$[30,40[$	$[40,50[$	$[50,60[$	$[60,70[$	هامشي Y
$[10,14[$	1	2	2	0	0	5
$[14,18[$	3	8	10	4	0	25
$[18,22[$	1	7	12	11	4	35
$[22,26[$	0	3	6	5	1	15
هامشي X	5	20	30	20	5	80

- حساب الانحراف المعياري من جدول التوزيع الهامشي لـ X

X	$[20,30[$	$[30,40[$	$[40,50[$	$[50,60[$	$[60,70[$	المجموع
n_i	5	20	30	20	5	80
C_i	25	35	45	55	65	
$n_{i.}C_i$	125	700	1350	1100	325	3600
$(C_i - \overline{X})$	-20	-10	0	10	20	
$(C_i - \overline{X})^2$	400	100	0	100	400	
$n_{i.}(C_i - \overline{X})^2$	2000	2000	0	2000	2000	8000

نلاحظ أن

$$\overline{X} = \frac{\sum_{i=1}^{5} C_i n_{i.}}{N} = \frac{3600}{80} = 45$$

وبالتالي

$$\sigma_X = \sqrt{\frac{1}{N}\sum_{i=1}^{5} n_{i.}\left(C_i - \overline{X}\right)^2} = \sqrt{\frac{8000}{80}} = 10$$

- حساب الانحراف المعياري من جدول التوزيع الهامشي لـ Y

Y	$[10,14[$	$[14,18[$	$[18,22[$	$[22,26[$	المجموع
$n_{.j}$	5	25	35	15	80
C_j	12	16	20	24	
$n_{.j}C_j$	60	400	700	360	1520
$\left(C_j - \overline{Y}\right)$	-7	-3	1	5	
$\left(C_j - \overline{Y}\right)^2$	49	9	1	25	
$n_{.j}\left(C_j - \overline{Y}\right)^2$	245	225	35	375	880

نلاحظ أن

$$\overline{Y} = \frac{1520}{80} = 19$$

وبالتالي

$$\sigma_Y = \sqrt{\frac{1}{N}\sum_{i=1}^{5} n_{.j}\left(C_j - \overline{Y}\right)^2} = \sqrt{\frac{880}{80}} = \sqrt{11} = 3,316$$

- حساب التغاير ومعامل الارتباط

$C_j-\overline{Y}$ \ $C_i-\overline{X}$	-20	-10	0	10	20	المجموع
-7	1(140)	2(70)	2(0)	0	0	280
-3	3(60)	8(30)	10(0)	4(-30)	0	300
1	1(-20)	7(-10)	12(0)	1(10)	4(20)	100
5	0	3(-50)	6(0)	5(50)	1(100)	200
المجموع	300	160	0	240	180	880

وبالتالي يكون:

$$Cov(X,Y)=\frac{1}{80}\sum_{i=1}^{5}\sum_{j=1}^{4}n_{ij}\left(C_i-\overline{X}\right)\left(C_j-\overline{Y}\right)=\frac{880}{80}=11$$

ومنه معمل الارتباط هو:

$$\rho=\frac{Cov(X,Y)}{\sigma_X\sigma_Y}=\frac{11}{10\sqrt{11}}=0,33$$

3.4. مستقيم الانحدار (التسوية الخطية)

لنفرض أنه لدينا مجتمع إحصائي مكون من N فردا وأننا نريد دراسته حسب صفتين في آن واحد. هدفنا هو تحديد وتحليل الارتباط بين هاتين الصفتين.

نملك على سبيل المثال مجموعة تحتوي على 10 أشخاص، وقسنا لكل واحد منهم طوله X ووزنه Y. وكان الغرض هو دراسة العلاقة بين هذين المتغيرين.

يمكننا ترتيب المعطيات في جدول كما يلي:

الوزن (كلغ)	الطول (متر)	الأفراد
58.2	1.55	رقم 1
58.1	1.60	رقم 2
61.3	1.62	رقم 3
65.3	1.64	رقم 4
69.5	1.65	رقم 5
69.7	1.70	رقم 6
70.3	1.72	رقم 7
75.4	1.73	رقم 8
74.2	1.76	رقم 9
82.0	1.78	رقم 10

في مرحلة أولى يزودنا حساب التباين المشترك ومعامل الارتباط الخطي بمعلومات حول العلاقة بين المتغيرين X و Y وحول طبيعتها.

إذا كان معامل الارتباط الخطي مجاورا لـ (1+) أو (1-)، فإن العلاقة بين X و Y تكون خطية، أي أنه بإمكاننا صياغتها في شكل معادلة مستقيم من الشكل:

$$Y = aX + b$$

وللوصول إلى صياغة نهائية للعلاقة بين X و Y فإن المطلوب هو تحديد معاملي المعادلة a و b، ويمكن حسابهما كما يلي:

$$a = \frac{Cov(X,Y)}{Var(X)}$$

و

$$b = \overline{Y} - a\overline{X}$$

عند تمثيلنا لهذه المعطيات في معلم، فإننا نحصل على سحابة نقاط.

- يسمى المستقيم الذي معادلته $Y = aX + b$ بمستقيم الانحدار أو مستقيم التسوية الخطية.

- يسمى a بمعامل الانحدار.

مثال تطبيقي:

نقوم بإيجاد معادلة مستقيم التسوية الخطية للمثال السابق كما يلي:

	X_i	Y_i	X_i^2	Y_i^2
	1.55	58.2	2.4025	90.21
	1.60	58.1	2.56	92.96
	1.62	61.3	2.6244	99.306
	1.64	65.3	2.6869	107.092
	1.65	69.5	2.7225	114.675
	1.70	69.7	2.89	118.49
	1.72	70.3	2.9584	120.916
	1.73	75.4	2.9929	130.442
	1.76	74.2	3.0976	130.592
	1.78	82.0	3.1684	145.96
$\sum$	16.75	684	28.1063	1150.643
$\sum / N$	1.675	68.4	2.81063	115.0643

لدينا

$$\overline{Y} = 68,4 \quad \text{و} \quad \overline{X} = 1,675$$

وبالتالي يكون

$$Var(X) = 2,81063 - (1,675)^2 = 0,005$$

$$Cov(X,Y) = 115,0643 - (1,675)(68,4) = 0,4943$$

وبالتالي فإن

$$a = \frac{Cov(X,Y)}{Var(X)} = \frac{0,4943}{0,005} = 98,86$$

$$b = \overline{Y} - a\overline{X} = 68,4 - (98,86)(1,675) = -97,19$$

ومنه معادلة مستقيم الانحدار هي

$$Y = 98,86X - 97,19$$

نتائج هامة:

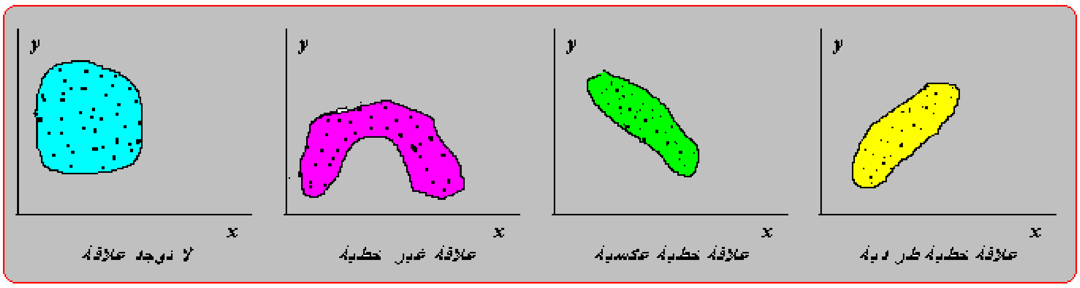
1) إن التباين المشترك يكون في قيمته المطلقة دوما أقل من جداء الانحرافين المعياريين أو مساويا له، و نكتب

$$|Cov(X,Y)| \leq \sigma_X \sigma_Y$$

2) كتابة a بدلالة ρ كما يلي

$$a = \rho \frac{\sigma_Y}{\sigma_X}$$

3) عند جمع البيانات للثنائية (X,Y) وتمثيلها بيانيا، فإن العلاقة بينها تأخذ أشكالا مختلفة على النحو التالي



الباب الثاني

مدخل إلى نظرية الاحتمالات

الفصل الأول

التحليل التوفيقي وأنواع السحب

في هذا الفصل، نذكر ببعض المفاهيم التي سوف نستخدمها في هذا الجزء من البرنامج، نبدأ قبل كل شيء بالتعريف ببعض المفاهيم الأساسية في التحليل التوفيقي كالمبدأ الأساسي في التحليل التوفيقي، الترتيبات، التبديلات وأخيرا التوفيقات ثم نعرض بعدها أنواع السحب.

الأهداف التعليمية للفصل:

يتعرف الطالب في هذا الفصل على:

- ✓ المبدأ الأساسي في التحليل التوفيقي
- ✓ الترتيبة بتكرار ودون تكرار
- ✓ التبديلة بتكرار ودون تكرار
- ✓ التوفيقات
- ✓ أنواع السحب.

1.1. التحليل التوفيقي

لتكن Ω مجموعة أساسية مكونة من n عنصر.

1.1.1. المبدأ الأساسي في التحليل التوفيقي

إذا استطعنا تمثيل عملية ما بـ n_1 طريقة مختلفة، ثم نجد بعدها أن عملية ثانية يمكن تشكيلها بـ n_2 طريقة مختلفة، وهكذا على التوالي، عندئذ يكون عدد الطرق المختلفة التي تسمح بإجراء العمليات المشار إليها بإتباع الترتيب هو الجداء:

$$n_1 \times n_2 \times ....$$

مثال:

تتضمن لوحة ترقيم سيارة حرفين لاتينيين مختلفين متبوعين بثلاثة أرقام يختلف أولها عن 0.

نحسب عدد اللوحات المختلفة الممكن الحصول عليها بهده الطريقة.

- عدد الأحرف اللاتينية هو 26 إذن هناك 26 طريقة لاختيار الحرف الأول أي أن $n_1 = 26$ ،
- توجد 25 طريقة للحصول على الحرف الثاني (لأن الحرف الثاني لا يمكن اختياره مطابقا للأول)، أي أن $n_2 = 25$ ،
- توجد 9 طرق مختلفة للحصول على الرقم الأول، أي أن $n_3 = 9$ ،
- توجد 10 طرق مختلفة للحصول على الرقم الثاني، أي أن $n_4 = 10$ ،
- هناك أيضا 10 طرق مختلفة للحصول على الرقم الثالث، أي أن $n_5 = 10$.

بتطبيق المبدأ الأساسي للتحليل التوفيقي نجد أن:

$$n_1 \times n_2 \times n_3 \times n_4 \times n_5 = 26 \times 25 \times 9 \times 10 \times 10 = 585000$$

إذن يمكننا الحصول على 585000 لوحة ترقيم مختلفة.

2.1.1. الترتيبات (Arrangements)

1) ترتيبه بتكرار (Arrangement avec répétition)

تعريف 1.1.1.

نسمي ترتيبه بتكرار كل تشكيلة مرتبة وبتكرار لـ p عنصر مأخوذة من n عنصر، وتعطى بالعلاقة التالية:

$$A_n^p = n^p$$

مثال:

في بلد ما، رقم الهاتف مكون من 5 أرقام، أوله يبدأ بالصفر والرقم الثاني يكون من 1 إلى 5، أما الأرقام الباقية فهي حرة من أي شرط.

نريد حساب كم رقم هاتف مختلف يمكن تشكيله في هذا البلد.

لدينا

$$\Omega = \{0,1,2,3,4,5,6,7,8,9\}$$

- توجد طريقة واحدة للحصول على الرقم الأول، أي أن $n_1 = 1$.
- توجد 5 طرق مختلفة للحصول على الرقم الثاني، أي أن $n_2 = 5$،
- توجد A_{10}^3 طرق مختلفة للحصول على الرقم الثالث، أي أن $n_3 = A_{10}^3 = 10^3 = 1000$،

وبتطبيق المبدأ الأساسي للتحليل التوفيقي نجد أن:

$$n_1 \times n_2 \times n_3 = 1 \times 5 \times A_{10}^3 = 1 \times 5 \times 1000 = 5000$$

وبالتالي يمكن تشكيل 5000 رقم هاتف.

2) ترتيبه دون تكرار (Arrangement sans répétition)

تعريف 2.1.1.

نسمي ترتيبه دون تكرار كل تشكيلة مرتبة وبدون تكرار لـ p عنصر مأخوذة من n عنصر، وتعطى بالعلاقة التالية:

$$A_n^p = \frac{n!}{(n-p)!}$$

ملاحظات:

$$n! = \prod_{i=1}^{n} i = 1 \times 2 \times \ldots \times n$$

$$0! = 1$$

مثال:

نستخدم نفس المثال السابق ولكن في هذه المرة دون تكرار بالنسبة للثلاثة أرقام الأخيرة.

$$n_1 \times n_2 \times n_3 = 1 \times 5 \times A_8^3 = 1 \times 5 \times \frac{8!}{(8-3)!} = 1 \times 5 \times 336 = 1680$$

وبالتالي يمكن تشكيل 1680 رقم هاتف.

3.1.1. التبديلات (Permutations)

1) تبديله دون تكرار (Permutationsans répétition)

تعريف 3.1.1.

نسمي تبديله دون تكرار، كل تشكيلة لـ n عنصر مأخوذة من n عنصر، وتعطى بالعلاقة التالية:

$$P_n = A_n^n = \frac{n!}{(n-n)!} = n!$$

ومنه

$$P_n = n!$$

2) تبديله بتكرار (Permutation avec répétition)

تعريف 4.1.1.

نسمي تبديله بتكرار، كل تشكيلة لـ n عنصر من بينها n_1 عنصرا متشابها و n_2 عنصرا متشابها و ... و n_k عنصرا متشابها، وتعطى بالعلاقة التالية:

$$P_n = \frac{n!}{n_1!\times n_2!\times\ldots\times n_k!}$$

مثال:

نرمي على التوالي زهرة نرد 12 مرة.

- نحسب عدد النتائج الممكنة أين كل وجه يظهر مرتين.

$$P_{12} = \frac{12!}{2!\times 2!\times 2!\times 2!\times 2!\times 2!}$$

- نحسب عدد النتائج الممكنة أين الرقم 1 يظهر 5 مرات، الرقم 2 يظهر 3 مرات، الرقم 3 يظهر 3 مرات والرقم 4 يظهر مرة واحدة.

$$P_{12} = \frac{12!}{5!\times 3!\times 3!\times 1!\times 0!\times 0!}$$

4.1.1. التوفيقات (Combinisons)

تعريف 5.1.1.

نسمي توفيقه، كل تشكيلة غير مرتبة لـ p عنصر مأخوذة من n عنصر. وتعطى بالعلاقة التالية:

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

ملاحظة:

$$A_n^p = p!C_n^p$$

خواص:

1) $C_n^n = C_n^0 = 1$.

2) $C_{n+1}^p = C_n^{p-1} + C_n^p$.

3) $(a+b)^n = \sum_{p=0}^{n} C_n^p a^p b^{n-p}$ (دستور ثنائي الحد لنيوتن).

مثال 1:

نسحب 5 أوراق من أوراق اللعب ذات 32 ورقة.

عدد النتائج الممكنة في هذه الحالة هو:

$$C_{32}^5 = \frac{32!}{5!(32-5)!}$$

مثال 2:

نريد تشكيل لجنة مكونة من 3 أفراد مأخوذة من 20 شخص.

عدد اللجان التي يمكن تشكيلها هو:

$$C_{20}^3 = \frac{20!}{3!(20-3)!} = 1140$$

مثال 3:

مجموعة مكونة من 5 رجال و 7 نساء.

عدد اللجان مختلفة التي يمكن تشكيلها من 2 رجال و 3 نساء هو:

$$C_5^2 \times C_7^3 = \frac{5!}{2!(5-2)!} \times \frac{7!}{3!(7-3)!} = 350$$

2.1. أنواع السحب

1.2.1. السحب بالإرجاع

يتم فيه اختيار عنصر من مجموعة ثم إعادته إليها، مما يعني أنه يمكن سحبه مرة أخرى في المرة التالية.

عدد الطرق الممكنة لسحب p عنصر مأخوذة من n عنصر مع الإرجاع لكل عنصر بعد السحب قبل إجراء السحب الموالي هو n^p.

مثال:

لدينا صندوق به 5 كرات $\{a,b,c,d,e\}$ ، نسحب كرتين على التوالي مع الإرجاع.

عدد الطرق الممكنة لهذا السحب هو:

$$n^p = 5^2 = 25.$$

2.2.1. السحب دون إرجاع

يتم فيه اختيار عنصر من مجموعة ثم استبعاده، فلا يمكن سحبه مرة أخرى. على سبيل المثال، إذا كان لدينا صندوق يحتوي على 5 كرات ملونة، عند سحب كرة وعدم إرجاعها، يتناقص عدد الكرات المتاحة للسحب التالي.

ونميز حالتين:

- **مع مراعاة الترتيب:** عدد الطرق الممكنة هو A_n^p.
- **دون مراعاة الترتيب:** عدد الطرق الممكنة هو C_n^p.

الفصل الثاني

التجارب العشوائية والفضاء الاحتمالي

في هذا الفصل، نذكر ببعض المفاهيم الأساسية لنظرية الاحتمال وخواصها وقوانينها، الحدث وأنواع الأحداث والتمييز بين قاعدة الجمع وقاعدة الجداء والإحداث المستقلة والمتنافية، بعدئذ نمر إلى التعريف بالاحتمال الشرطي ونظرية بايز.

الأهداف التعليمية للفصل:

يتعرف الطالب في هذا الفصل على:

- ✓ المفاهيم الأساسية لنظرية الاحتمال
- ✓ الأحداث البسيطة والمركبة
- ✓ احتمال الحدث
- ✓ الاحتمال الشرطي والأحداث المستقلة
- ✓ قانون الاحتمال الكلي وقاعدة بايز.

1.2. تذكرة وتعاريف هامة

نذكر أولا ببعض المفاهيم الأساسية في نظرية القياس.

تعريف 1.1.2. (مجموعة الأجزاء)

من أجل كل مجموعة Ω، نرمز بـ $\mathcal{P}(\Omega)$ بمجموعة أجزاء هذه المجموعة.

مثال:

لتكن $\Omega=\{1,2,3\}$، مجموعة أجزاء Ω هي:

$$\mathcal{P}(\Omega)=\{\phi,\Omega,\{1\},\{2\},\{3\},\{1,2\},\{1,3\},\{2,3\}\}$$

تعريف 2.1.2. (أصلي مجموعة)

نسمي أصلي مجموعة منتهية Ω، عدد عناصر هذه المجموعة. ونكتب $Card\ (\Omega)$.

مثال:

لتكن $\Omega=\{0,1,...,10\}$، يكون لدينا:

$$Card\ (\Omega)=11$$

تعريف 3.1.2. (العشيرة)

نسمي $\mathcal{F}\subset\mathcal{P}(\Omega)$ عشيرة أو ($\sigma-$جبر) إذا كان:

1) $\Omega\in\mathcal{F}$.

2) إذا كان $\forall i\in IN$ ، $A_i\in\mathcal{F}$ فإن $\underset{i\in IN}{\cup}A_i\in\mathcal{F}$.

3) إذا كان $A\in\mathcal{F}$ إذن $A^c\in\mathcal{F}$.

- نسمي مجموعة قابلة للقياس، كل مجموعة A ، حيث A عنصر في العشيرة $\mathcal{F}$.
- كل مجموعة مزودة بعشيرة تسمى فضاء قابل للقياس، ونكتب $(\Omega,\mathcal{F})$.

تعريف 4.1.2. (العشيرة المولدة)

لتكن $\mathcal{M}$ جزء من $\mathcal{P}(\Omega)$. نسمي عشيرة مولدة بـ $\mathcal{M}$، تقاطع كل العشائر التي تحتوي على $\mathcal{M}$، و نكتب $\sigma(\mathcal{M})$.

تعريف 5.1.2. (العشيرة البوريلية)

ليكن (Ω,τ) فضاء طبولوجي، نسمي عشيرة بوريلية على Ω نسبة إلى τ و نرمز لها بـ $\mathcal{B}_\tau(\Omega)$، العشيرة المولدة بمفتوحات Ω.

تعريف 6.1.2. (قابلية القياس)

نقول أن التطبيق $f:(\Omega,\mathcal{A})\to(E,\mathcal{F})$ قابل للقياس، إذا كان:

$$\forall B\in\mathcal{F},\quad f^{-1}(B)\in\mathcal{A}$$

حيث أن $(\Omega,\mathcal{A})$ و $(E,\mathcal{F})$ فضاءان قابلان للقياس.

تعريف 7.1.2. (القياس)

ليكن $(\Omega,\mathcal{F})$ فضاء قابل للقياس، نسمي التطبيق $\mu:\mathcal{F}\to[0,+\infty]$، قياس إذا وفقط إذا كان:

1) $\mu(\phi)=0$.

2) إذا كانت $(A_n)_{n\in IN^*}$ متتالية قابلة للعد لمجموعات من $\mathcal{F}$، منفصلة مثنى مثنى فإن:

$$\mu\left(\bigcup_{n=1}^{+\infty}A_n\right)=\sum_{n=1}^{+\infty}\mu(A_n)$$

تعريف 8.1.2. (الاحتمال)

نسمي احتمال على الفضاء $(\Omega,\mathcal{F})$، كل تطبيق IP حيث: $IP:\mathcal{F}\to[0,1]$ يحقق

1) $IP(\Omega)=1$.

2) من أجل كل متتالية $(A_n)_{n\in IN^*}$ حيث $A_n\in\mathcal{F}$، منفصلة مثنى مثنى، لدينا:

$$IP\left(\bigcup_{n=1}^{+\infty}A_n\right)=\sum_{n=1}^{+\infty}IP(A_n)$$

2.2. مفاهيم أساسية في الاحتمالات

1.2.2. التجربة العشوائية

تعريف 1.2.2. نسمي تجربة عشوائية كل تجربة لا يمكن التنبؤ بنتائجها مسبقا قبل إجرائها بدقة.

مثال:

تجربة رمي زهرة نرد، تجربة رمي قطعة نقدية،...

2.2.2. فضاء العينة

تعريف 2.2.2. نسمي Ω المجموعة الأساسية أو مجموعة النتائج الممكنة أو مجموعة الإمكانيات.

مثال:

في تجربة رمي زهرة نرد، مجموعة الإمكانيات هي $\Omega=\{1,2,3,4,5,6\}$.

في تجربة رمي قطعة نقدية، مجموعة الإمكانيات هي $\Omega=\{F,P\}$.

3.2.2. الأحداث

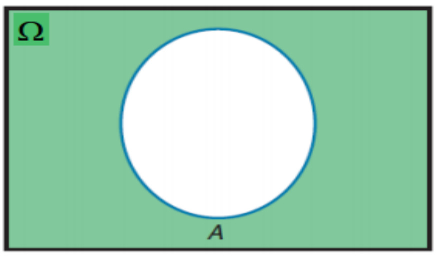
1) نسمي **حدث** أو حادثة A كل جزء من مجموعة إمكانيات التجربة العشوائية.

مثال

في تجربة رمي زهرة نرد مرة واحدة، المجموعات التالية تشكل أحداث لأنها مجموعات جزئية من فضاء العينة Ω:

- ✓ ظهور عدد فردي، أي أن $A=\{1,3,5\}$.
- ✓ ظهور عدد أكبر تماما من 2، أي أن $B=\{3,4,5,6\}$.

2) **الحدث المعاكس** أو المتمم للحدث A هو الحدث المكون من جميع عناصر فضاء العينة التي لا تنتمي إلى A، ونكتب $\overline{A}=A^c=\Omega-A$.



مثال:

في تجربة رمي زهرة نرد، الحدث A هو حدث الحصول على وجه يحمل رقم زوجي أي أن $A=\{2,4,6\}$.

- ✓ الحدث المعاكس للحدث A هو $\overline{A}=\{1,3,5\}$.

3) الحدث Ω يسمى **الحدث الأكيد** وهو الحدث المكون من مجموعة النتائج الممكنة للتجربة.

4) الحدث ϕ يسمى **الحدث المستحيل** وهو الحدث الذي لا يحتوي على أية نتيجة، كظهور عدد سالب في تجربة رمي زهرة نرد.

5) **الحدث الابتدائي** (الأولي) هو كل مجموعة مكونة من عنصر واحد فقط.

مثال1:

الحوادث الأولية في تجربة رمي قطعة نقدية هي $\{F\}$ ، $\{P\}$.

مثال2:

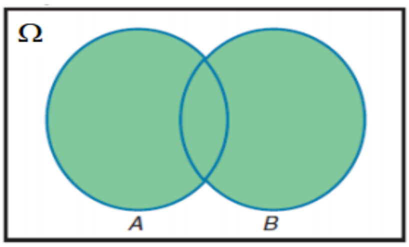
الحوادث الأولية في تجربة رمي قطعتين نقديتين هي $\{FP\}$ ، $\{PF\}$ ، $\{PP\}$ ، $\{FF\}$.

6) **الحدث المركب** هو كل مجموعة مكونة من عنصرين أو أكثر.

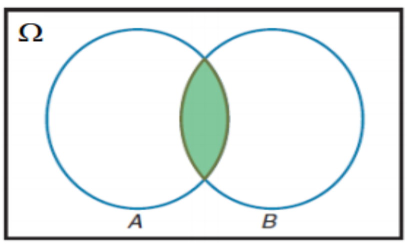
مثال:

في تجربة رمي زهرة نرد، نسمي حادثة الحصول على وجه علوي يحمل رقما فرديا بالحدث المركب أي أن $A=\{1,3,5\}$

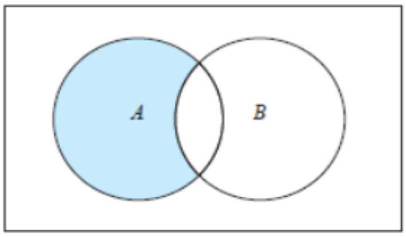
7) الحدث $A\cup B$ هو الحدث المكون من جميع العناصر التي تنتمي إلى A أو تنتمي إلى B أو تنتمي إلى كليهما (معناه وقوع A أو وقوع B أو كليهما).



8) الحدث $A\cap B$ هو الحدث المكون من جميع العناصر التي تنتمي إلى A وتنتمي إلى B في آن واحد (معناه وقوع A ووقوع B معا).



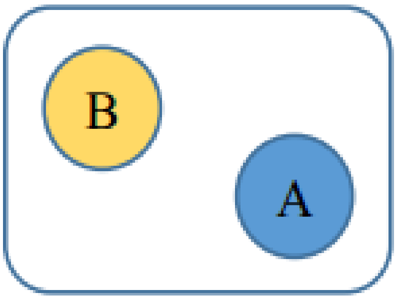
9) الحدث $A-B$ هو الحدث المكون من جميع العناصر الموجودة في A غير الموجودة في B (معناه وقوع A وعدم وقوع B).



ملاحظة: نلاحظ أن

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

10) **الحوادث المتنافية** (المنفصلة) هي الحوادث التي لا يمكن أن تتحقق معا في آن واحد وبالتالي يعطي تقاطعها مجموعة خالية، ونكتب $A\cap B=\phi$.



3.2. فضاء الاحتمال

تعريف 1.3.2.

نسمي فضاء احتمال، الثلاثية $(\Omega,\mathcal{F},IP)$ حيث:

1) Ω : مجموعة.
2) $\mathcal{F}$: عشيرة أحداث، وعناصر $\mathcal{F}$ تسمى أحداث.
3) IP : قياس موجب على العشيرة $\mathcal{F}$، حيث $IP(\Omega)=1$ و $0\leq IP\leq 1$.

4.2. خواص الاحتمال

كل احتمال IP معرف على الفضاء $(\Omega,\mathcal{F})$ يحقق الخواص التالية:

1) $IP(\Omega)=1$. (من تعريف الاحتمال)

2) $IP(\overline{A})=1-IP(A)$

لدينا $\Omega=A\cup\overline{A}$ مع الاتحاد المنفصل، يكون لدينا إذن

$$1=IP(\Omega)=IP\left(A\cup\overline{A}\right)=IP(A)+IP\left(\overline{A}\right)$$

ومنه

$$IP\left(\overline{A}\right)=1-IP(A)$$

3) $IP(\phi)=0$.

لدينا $\overline{\Omega}=\phi$ ، إذن

$$IP(\phi)=IP\left(\overline{\Omega}\right)=1-IP(\Omega)=1-1=0$$

ومنه

$$IP(\phi)=0$$

4) حالات خاصة للخاصية الثانية من تعريف الاحتمال

- ✓ إذا كان $A\cap B=\phi$ ، فإن $IP(A\cup B)=IP(A)+IP(B)$
- ✓ إذا كانت $(A_i)_{1\le i\le n}$ منفصلة مثنى مثنى، فإن

$$IP\left(\bigcup_{i=1}^{n}A_i\right)=\sum_{i=1}^{n}IP(A_i)$$

5) $IP(A\cup B)=IP(A)+IP(B)-IP(A\cap B)$ ، حيث A و B حدثان كيفيان.

لدينا $A\cup B=(A\cap B)\cup(B-A)\cup(A-B)$ حيث أن المجموعات منفصلة مثنى مثنى.

يكون لدينا إذن

$$IP(A\cup B)=IP(A\cap B)+IP(B-A)+IP(A-B) \qquad (1)$$

ولدينا $A=(A\cap B)\cup(A-B)$ مع اتحاد المجموعات المنفصلة، إذن

$$IP(A)=IP(A\cap B)+IP(A-B)$$

أي أن

$$IP(A-B)=IP(A)-IP(A\cap B)$$

ولدينا $B=(A\cap B)\cup(B-A)$ مع اتحاد المجموعات المنفصلة، إذن

$$IP(B)=IP(A\cap B)+IP(B-A)$$

أي أن

$$IP(B-A)=IP(B)-IP(A\cap B)$$

وبالتعويض في (1) نجد أن

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

6) $\forall \mathrm{A},\mathrm{B}\in \mathcal{F},\ \mathrm{A}\subset \mathrm{B}\Rightarrow IP(A)\leq IP(B)$

لدينا $B=A\cup(B-A)$ أين الاتحاد منفصل، ويكون لدينا إذن

$$IP(B)=IP(A)+IP(B-A)\geq IP(A)$$

ومنه

$$IP(B)\geq IP(A)$$

7) الاستمرار المتزايد و المتناقص

لتكن متتالية الحوادث $B_1,B_2,\ldots,B_n,\ldots$ من $\mathcal{F}$ ولتكن متزايدة (أي أن من أجل كل $n\geq 1$، لدينا $B_n\subset B_{n+1}$) أو متناقصة (أي أن من أجل كل $n\geq 1$، لدينا $B_n\supset B_{n+1}$)، إذن

✓ في حالة التزايد

$$\lim_{n\to+\infty} IP(B_n)=IP\left(\underset{n\geq 1}{\cup} B_n\right)$$

لدينا $\overset{n}{\underset{k=1}{\cup}} B_k=B_n$ ، إذن

$$\lim_{n\to+\infty} IP(B_n)=\lim_{n\to+\infty} IP\left(\overset{n}{\underset{k=1}{\cup}} B_k\right)=IP\left(\overset{+\infty}{\underset{k=1}{\cup}} B_k\right)$$

وحسب بديهية الجمع القابلة للعد لدينا

$$\sum_{k=1}^{n} IP(B_k)=IP\left(\underset{k\geq 1}{\cup} B_k\right)$$

✓ في حالة التناقص

$$\lim_{n\to+\infty} IP(B_n)=IP\left(\bigcap_{n\geq 1} B_n\right)$$

لدينا

$$\lim_{n\to+\infty} IP(B_n)=1-\lim IP(B_n^c)=1-IP\left(\bigcup_{n\geq 1} B_n^c\right)=1-\left(1-IP\left(\bigcap_{n\geq 1} B_n\right)\right)=IP\left(\bigcap_{n\geq 1} B_n\right)$$

أو باستخدام الطريقة الأولى كما يلي

لدينا $\bigcap_{k=1}^{n} B_k = B_n$ ، إذن

$$\lim_{n\to+\infty} IP(B_n)=\lim_{n\to+\infty} IP\left(\bigcap_{k=1}^{n} B_k\right)=IP\left(\bigcap_{k=1}^{+\infty} B_k\right)$$

8) الجمعية الجزئية القابلة للعد

لتكن متتالية الحوادث $B_1,B_2,...,B_n,...$ من $\mathcal{F}$. إذن السلسلة $\sum_{k=1}^{\infty} IP(B_k)$ إما متقاربة أو متباعدة و في هذه الحالة يكون

$$\sum_{k=1}^{\infty} IP(B_k) > IP\left(\bigcup_{n\geq 1} B_n\right)$$

متتالية الحوادث المعرفة بـ $C_n=B_1\cup B_2\cup...\cup B_n$ متزايدة، نستطيع تطبيق الخاصية

$$IP(B_1\cup B_2\cup...\cup B_n)=IP(B_1)+IP(B_2)+...+IP(B_n)$$

حيث $B_1,B_2,...,B_n,...$ منفصلة مثنى مثنى، فنجد أن

$$IP\left(\bigcup_{n\geq 1} B_n\right)=\lim_{n\to+\infty} IP(C_n)\leq \lim_{n\to+\infty}\left(IP(B_1)+IP(B_2)+...+IP(B_n)\right)=\sum_{k=1}^{\infty} IP(B_k)$$

5.2. تساوي الاحتمال

إن أبسط الحالات في حساب الاحتمال هي تلك التي تتعلق بتجارب تتصف بعدد منته من الأحداث الابتدائية وتكون الأحداث الابتدائية متساوية الاحتمال، أي أن $IP_1 = IP_2 = \ldots = IP_n$ فإن

$$\forall\, i, \quad IP_i = IP = \frac{1}{N} = \frac{1}{Card(\Omega)}$$

مثال:

في تجربة رمي زهرة نرد، احتمال الحوادث الأولية هو:

$$IP(1) = IP(2) = IP(3) = IP(4) = IP(5) = IP(6) = \frac{1}{6}$$

نقول في هذه الحالة أن الاحتمال متساوي التوزيع.

ومن أجل حدث A (مركب أو ابتدائي) حيث أن $A \in \mathcal{P}(\Omega)$ ، فإن

$$IP(A) = \frac{Card(A)}{Card(\Omega)} = \frac{n}{N}$$

مثال1

في تجربة رمي زهرة نرد، A حادثة الحصول على وجه يحمل رقم مضاعف للعدد 3، أي أن $A = \{3,6\}$ ، ومنه احتمال الحادثة A هو:

$$IP(A) = \frac{Card(A)}{Card(\Omega)} = \frac{2}{6} = \frac{1}{3}$$

مثال 2

يحتوي صندوق على 4 كرات بيضاء و 5 كرات حمراء و 6 كرات صفراء، تسحب عشوائيا ثلاث كرات من الصندوق معا. لتكن الأحداث التالية:

a) A الحصول على ثلاث كرات حمراء،

b) B الحصول على ثلاث كرات لها نفس اللون،

c) C الحصول على ثلاث كرات مختلفة الألوان،

d) D الحصول على كرة صفراء واحدة على الأقل.

أحسب احتمال كل من الأحداث A, B, C, D .

الحل

لدينا $N = C_{15}^3 = 455$

$$IP(A) = \frac{C_6^3}{C_{15}^3} = \frac{20}{455} \quad (a$$

$$IP(B) = \frac{C_4^3 + C_5^3 + C_6^3}{C_{15}^3} = \frac{34}{455} \quad (b$$

$$IP(C) = \frac{C_4^1 + C_5^1 + C_6^1}{C_{15}^3} = \frac{120}{455} \quad (c$$

$$IP(D) = \frac{C_6^1 . C_9^2 + C_6^2 . C_9^1 + C_6^3}{C_{15}^3} = \frac{371}{455} \quad (d$$

6.2. الاحتمال الشرطي

من المسائل المهمة في حساب الاحتمال دراسة العلاقات الاحتمالية ما بين الأحداث، فإذا كان A و B حدثين متعلقين بتجربة معينة فإن وقوع أحد هذين الحدثين قد يؤثر على احتمال وقوع الحدث الآخر.

تعريف 1.6.2.

ليكن $(\Omega, \mathcal{F}, IP)$ فضاء احتمال، وليكن A و B حدثين معرفين على نفس الفضاء، حيث أن $IP(B) \neq 0$.

الاحتمال الشرطي للحدث A علما أن الحدث B قد وقع، ونرمز له بالرمز $IP(A/B)$ أو $IP_B(A)$، يعرف كما يلي:

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

مثال:

احتمال اقلاع (مغادرة) رحلة طيران نظامية في الوقت المحدد لها هو $IP(D)=0.83$ ، واحتمال أن تصل في الوقت المحدد لها هو $IP(A)=0.82$ ، كما أن احتمال المغادرة والوصول في الوقت المحدد لها هو $IP(A\cap D)=0.78$. أوجد احتمال أن:

a) تصل الطائرة في الوقت المحدد علما أنها أقلعت في الوقت المحدد.
b) تقلع الطائرة في الوقت المحدد علما أنها وصلت في الوقت المحدد.
c) تصل الطائرة في الوقت المحدد علما أنها لم تقلع في الوقت المحدد.

الحل:

$$IP(A/D)=\frac{IP(A\cap D)}{IP(D)}=\frac{0.78}{0.83}=0.94 \quad (a$$

$$IP(D/A)=\frac{IP(D\cap A)}{IP(A)}=\frac{0.78}{0.82}=0.95 \quad (b$$

$$IP(A/\overline{D})=\frac{IP(A\cap \overline{D})}{IP(\overline{D})}=\frac{IP(A)-IP(A\cap D)}{1-IP(D)}=\frac{0.82-0.78}{1-0.83}=\frac{0.04}{0.17}=0.24 \quad (c$$

نتيجة:

نستنتج من تعريف الاحتمال الشرطي أن:

1) إذا كان $A\cap B=\phi$ ، فإن $IP(A\cup B/C)=IP(A/C)+IP(B/C)$

2) $IP(\overline{A}/B)=1-IP(A/B)$

3) $IP(A\cap B)=IP(A/B)IP(B)=IP(B/A)IP(A)$

مثال:

وعاء يحتوي على 12 كرة، 5 خضراء و 7 زرقاء، سحبنا كرتين (سحب دون إرجاع ودون مراعاة الترتيب).

- نحسب احتمال أن يكونا زرقاوين.

✓ نضع A_1 حدث ظهور كرة زرقاء في المرة الأولى.

✓ نضع A_2 حدث ظهور كرة زرقاء في المرة الثانية.

يصبح في هذه الحالة حدث ظهور كرتين زرقاوين هو الحدث $A_1 \cap A_2$

$$IP(A_1 \cap A_2) = IP(A_1).IP(A_2/A_1) = \frac{C_7^1}{C_{12}^1} \times \frac{C_6^1}{C_{11}^1} = \frac{7}{12}.\frac{6}{11} = \frac{C_7^2}{C_{12}^2} = \frac{7}{22}$$

7.2. قاعدة بايز (احتمال السبب)

لنفرض أن هناك عدد من الأسباب المعينة التي يؤدي وقوع أحدها إلى حدوث حادثة ما. وأن هذه الحادثة تقع إذا وقع أحد أسبابها. ولنفرض أننا نعلم مسبقا احتمال تحقق كل سبب من هذه الأسباب وكذلك نعلم الاحتمال الشرطي لهذه الحادثة عند تحقق كل سبب من أسبابها. إن نظرية بايز تعني بحساب احتمال أن يكون سببا محددا من الأسباب هو مصدر حدوث هذه الحادثة والتي نعلم مسبقا بحدوثها. وقبل استعراض نظرية بايز فإنه لابد من التطرق لما يسمى بقانون الاحتمال الكلي الذي يعني بحساب وقوع هذه الحادثة بغض النظر عن السبب.

1.7.2. الاحتمال الكلي

قضية 1.7.2. $[11]$

ليكن $(\Omega, \mathcal{F}, IP)$ فضاء احتمال. الأحداث $A_1, A_2, ..., A_n$ تجزئة لـ Ω، متنافية مثنى مثنى، حيث أن $\Omega = \bigcup_{i=1}^{n} A_i$، إذا كان B حدث حيث $(B \in \mathcal{F})$، فعندئذ يكون لدينا

$$IP(B) = \sum_{i=1}^{n} IP(B \cap A_i) = \sum_{i=1}^{n} IP(A_i).IP(B/A_i)$$

البرهان

لدينا

$$B = B \cap \Omega = B \cap \left(\bigcup_{i=1}^{n} A_i\right) = \bigcup_{i=1}^{n}(B \cap A_i)$$

ومن جهة أخرى

$$IP(B) = IP\left(\bigcup_{i=1}^{n}(B \cap A_i)\right) = \sum_{i=1}^{n} IP(B \cap A_i)$$

و

$$IP(B)=IP(A_i).IP(B/A_i)$$

مثال:

وعاء يحتوي الأول على 5 كرات، 3 حمراء و 2 زرقاء، ويحتوي الثاني على 2 حمراء و 8 زرقاء، ولهما نفس الأفضلية للسحب.

- نحسب احتمال سحب كرة حمراء.

✓ نضع H حدث ظهور كرة حمراء.

✓ نضع A_1 سحب كرة من الوعاء الأول.

✓ نضع A_2 سحب كرة من الوعاء الثاني.

ومنه يكون:

$$\begin{aligned}IP(H)&=\sum_{i=1}^{2}IP(A_i).IP(H/A_i)\\&=IP(A_1).IP(H/A_1)+IP(A_2).IP(H/A_2)\end{aligned}$$

وبالتالي ينتج أن:

$$\begin{aligned}IP(H)&=\frac{1}{2}.\frac{C_3^1}{C_5^1}+\frac{1}{2}.\frac{C_2^1}{C_{10}^1}\\&=\frac{1}{2}.\frac{3}{5}+\frac{1}{2}.\frac{2}{10}=\frac{2}{5}\end{aligned}$$

2.7.2. قاعدة بايز

نظرية 1.7.2. (نظرية بايز) [11]

إذا كانت $A_1, A_2,..., A_n$ تجزئة لـ Ω، وكان $(B\in\mathcal{F})$ في الفضاء $(\Omega,\mathcal{F},IP)$، فعندئذ يكون لدينا:
$\forall j\in\{1,2,.......n\}$

$$P(A_j/B)=\frac{P(A_j\cap B)}{P(B)}=\frac{P(A_j).P(B/A_j)}{\sum_{i=1}^{n}IP(A_i).IP(B/A_i)}$$

وهذا ينتج من تعريف الاحتمال الشرطي، ومن القضية السابقة، ويمكن أن ننظر إلى الأحداث $(A_i)_{i=1,n}$ على أنها تمثل دور الأسباب التي أدت إلى وقوع الحدث B ، وذلك أن أحدهما فقط يقع. فإذا وقع الحدث B ، فإن احتمال أن يكون A_j هو السبب، هو الاحتمال

$$P(A_j/B)=\frac{P(A_j).P(B/A_j)}{\sum_{i=1}^{n} IP(A_i).IP(B/A_i)}$$

مثال:

لتكن لدينا ثلاث آلات H_1, H_2, H_3 تنتج على الترتيب 40%، 35%، 25% من إنتاج مصنع ما، نفرض أن 2%، 4%، 5% من إنتاج هذه الآلات رديء. أخذنا عينة من هذا الإنتاج ووجدناها رديئة.

- نحسب احتمال أن تكون من إنتاج الآلة H_1.

حسب قاعدة بايز نضع A_i السلعة من إنتاج H_i حيث $1 \le i \le 3$ ، و B سحبنا سلعة رديئة.

المقصود إذن هو حساب $IP(A_1/B)$.

لدينا

$$IP(A_1/B)=\frac{IP(A_1).IP(B/A_1)}{IP(A_1).IP(B/A_1)+IP(A_2).IP(B/A_2)+IP(A_3).IP(B/A_3)}$$

حيث أن

$$IP(A_1)=\frac{25}{100}=0,25; \qquad IP(B/A_1)=\frac{2}{100}=0,02$$

$$IP(A_2)=\frac{35}{100}=0,35; \qquad IP(B/A_2)=\frac{4}{100}=0,04$$

$$IP(A_3)=\frac{40}{100}=0,4; \qquad IP(B/A_3)=\frac{5}{100}=0,05$$

إذن

$$IP(A_1/B)=\frac{0,25.0,02}{0,25.0,02+0,35.0,04+0,4.0,05}=\frac{5}{39}$$

8.2. الحدثان المستقلان

تعريف 1.8.2.

إذا كان $A,B\in\mathcal{F}$ حدثان وكان وقوع A لا يؤثر في احتمال وقوع B، فإن

$$IP(B/A)=IP(B)$$

فنقول عندئذ أن A و B حدثان مستقلان.

وبما أن

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

فإن

A و B مستقلان $\Leftrightarrow IP(A\cap B)=IP(A)\cdot IP(B)$

مثال:

رمى شخصان في اتجاه لوح واحد معلق في جدار، واحتمال إصابتهما للهدف هو $\frac{4}{10}$ للأول و $\frac{7}{10}$ للثاني.

- نحسب احتمال إصابة الهدف
  - ✓ نضع A_1 حادثة إصابة الأول الهدف، أي أن $IP(A_1)=\frac{4}{10}$.
  - ✓ نضع A_2 حادثة إصابة الثاني الهدف، أي أن $IP(A_2)=\frac{7}{10}$.

حدث إصابة الهدف هو $A_1\cup A_2$، أي أنه علينا حساب $IP(A_1\cup A_2)$

لدينا

$$IP(A_1\cup A_2)=IP(A_1)+IP(A_2)-IP(A_1\cap A_2)$$

نلاحظ أن الحدثان A_1 و A_2 مستقلان أي أن $IP(A_1\cap A_2)=IP(A_1).IP(A_2)$

وبالتالي ينتج

$$IP(A_1 \cup A_2) = IP(A_1) + IP(A_2) - IP(A_1).IP(A_2)$$

$$= \frac{4}{10} + \frac{7}{10} - \frac{28}{100} = \frac{82}{100} = \frac{41}{50}$$

الفصل الثالث

المتغيرات العشوائية

نتطرق في هذا الفصل إلى دراسة المتغيرات العشوائية وحيدة البعد وخصائصها من قانون احتمال وتابع توزيع وغيرها من الخصائص، كما سنعرض أشهر التوزيعات الاحتمالية المتقطعة منها والمستمرة التي يمكن أن يكون أي منها توزيعا احتماليا لعدد من المتغيرات العشوائية المتماثلة في سلوكها والمختلفة ببعض أو كل مميزاتها العددية أو بالثوابت التي تدخل في تشكيل قانون توزيعها الاحتمالي.

الأهداف التعليمية للفصل:

يتعرف الطالب في هذا الفصل على:

- ✓ المتغيرات العشوائية الحقيقية
- ✓ تابع التوزيع
- ✓ قوانين الاحتمال الشهيرة
- ✓ الأمل الرياضي والتباين
- ✓ التابع المولد للعزوم والتابع المميز.

1.3. المتغير العشوائي الحقيقي

تعريف 1.1.3.

ليكن $(\Omega,\mathcal{F},IP)$ فضاء احتمال. نسمي متغير عشوائي حقيقي معرف على الفضاء $(\Omega,\mathcal{F},IP)$ كل تطبيق قابل للقياس

$$X:(\Omega,\mathcal{F},IP)\rightarrow\left(IR,\mathcal{B}_{IR}\right)$$

تقسم المتغيرات العشوائية إلى نوعين أساسيين هما:

- **المتغير العشوائي المتقطع** أو النقطي هو الذي يأخذ عددا منتهيا أو قابلا للعد من القيم.
- **المتغير العشوائي المستمر** هو الذي يأخذ كل قيم مجال محدود أو غير محدود.

مثال:

نرمي قطعة نقدية متزنة مرتين متتاليتين، نجد أن $\Omega=\{FF,FP,PF,PP\}$

نعتبر أن X هو عدد مرات ظهور الوجه F في هذه التجربة، فنجد أن:

$$\begin{aligned}&X(PP)=0\\&X(FP)=X(PF)=1\\&X(FF)=2\end{aligned}$$

وبالتالي تكون قيم المتغير العشوائي كما يلي:

$$X(\Omega)=\{0,1,2\}$$

2.3. قانون احتمال المتغير العشوائي

تعريف 1.2.3.

ليكن X متغير عشوائي حقيقي، حيث $X:(\Omega,\mathcal{F},IP)\rightarrow\left(IR,\mathcal{B}_{IR}\right)$

نعرف التابع IP_X على العشيرة $\mathcal{B}_{IR}$ كما يلي:

$$\forall B\in\mathcal{B}_{IR},\quad IP_X(B)=IP(X\in B)=IP\Big(\omega\in\Omega;\quad X(\omega)\in B\Big)=IP\left(X^{-1}(B)\right)$$

مثال:

نلقي قطعة نقدية ثلاث مرات متتالية، وليكن X هو عدد مرات ظهور الوجه (F)، فيكون لدينا:

$$\Omega=\{FFF,FFP,FPF,PFF,PPF,PFP,FPP,PPP\}$$
$$\omega_1=FFF,\ \ \omega_2=FFP,\ \ \omega_3=FPF,\ \ \omega_4=PFF,$$
$$\omega_5=PPF,\ \ \omega_6=PFP,\ \ \omega_7=FPP,\ \ \omega_8=PPP.$$

$$\mathcal{F}=\mathcal{P}(\Omega) \quad \text{و} \quad IP(\omega_i)=\frac{1}{8}$$

$$\begin{aligned}
&X(\omega_1)=3\\
&X(\omega_2)=X(\omega_3)=X(\omega_4)=2\\
&X(\omega_5)=X(\omega_6)=X(\omega_7)=1\\
&X(\omega_8)=0
\end{aligned}$$

وبالتالي

$$\mathrm{X}(\Omega)=\{0,1,2,3\}$$

قانون احتمال المتغير العشوائي X :

- ✓ احتمال عدم ظهور أي وجه $IP(X=0)=IP(PPP)=\frac{1}{8}$
- ✓ احتمال ظهور وجه واحد $IP(X=1)=IP(PPF,PFP,FPP)=\frac{3}{8}$
- ✓ احتمال ظهور وجهين $IP(X=2)=IP(PFF,PFF,FPF)=\frac{3}{8}$
- ✓ احتمال ظهور ثلاث أوجه $IP(X=3)=IP(FFF)=\frac{1}{8}$

يعطى قانون الاحتمال جدوليا كما يلي:

X	0	1	2	3
IP_X	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

3.3. تابع توزيع المتغير العشوائي

تعريف 1.3.3. (تابع التوزيع)

نسمي تابع توزيع المتغير العشوائي الحقيقي X المعرف على فضاء الاحتمال $(\Omega,\mathcal{F},IP)$، التابع F_X المعرف على IR بـ:

$$\forall x\in IR, \qquad F_X(x)=IP(X\leq x)=IP(]-\infty,x])$$

قضية 1.3.3.(خواص تابع التوزيع)

1) F_X متزايدة على IR.

2) $\lim\limits_{t\to-\infty}F_X(t)=0$ و $\lim\limits_{t\to+\infty}F_X(t)=1$.

3) F_X مستمرة من اليمين.

4) F_X لديها نهاية من اليسار عند كل نقطة.

البرهان

1) إذا كان $x\leq y$، إذن $X\leq x\Rightarrow X\leq y$ و منه $IP(X\leq x)\leq IP(X\leq y)$.

2) بما أن F_X متزايدة،

$$\lim_{t\to-\infty}F_X(t)=\lim_{n\to+\infty}F_X(-n)=\lim_{n\to+\infty}IP_X(]-\infty,-n])$$
$$=IP_X\left(\bigcap_{n=1}^{+\infty}]-\infty,-n]\right)=IP_X(\phi)=0$$

نفس الشيء بالنسبة لـ

$$\lim_{t\to+\infty}F_X(t)=\lim_{n\to+\infty}F_X(n)=\lim_{n\to+\infty}IP_X(]-\infty,n])$$
$$=IP_X\left(\bigcup_{n=1}^{n}]-\infty,n]\right)=IP_X(IR)=1$$

3) بما أن F_X متزايدة،

$$\lim_{y\to x^+}F_X(y)=\lim_{n\to+\infty}F_X\left(x+\frac{1}{n}\right)=\lim_{n\to+\infty}IP_X\left(\left]-\infty,x+\frac{1}{n}\right]\right)$$
$$=IP_X\left(\bigcap_{n=1}^{+\infty}\left]-\infty,x+\frac{1}{n}\right]\right)=IP_X(]-\infty,x])=F_X(x).$$

4) بما أن F_X متزايدة و محدودة، فهي تملك نهاية من اليسار عند كل نقطة.

1.3.3. تابع التوزيع في حالة المتغير العشوائي المتقطع

ليكن X متغير عشوائي متقطع معرف على فضاء الاحتمال $(\Omega,\mathcal{F},IP)$، يمكن تعريف التابع F_X كما يلي:

$$F_X(x)=\begin{cases} 0 & ;\ -\infty < x < x_1 \\ IP(X=x_1) & ;\ x_1 \le x < x_2 \\ IP(X=x_1)+IP(X=x_2) & ;\ x_2 \le x < x_3 \\ . \\ . \\ . \\ IP(X=x_1)+IP(X=x_2)+\ldots+IP(X=x_n) & ;\ x \ge x_n \end{cases}$$

مثال:

في تجربة رمي زوج من النرد المتزن مرة واحدة. ليكن Y المتغير العشوائي الذي يدل على مجموع أرقام الوجهين الظاهرين. أوجد تابع التوزيع الاحتمالي للمتغير Y.

الحل:

من معطيات التجربة نجد أن قيم المتغير العشوائي هي:

$$Y(\Omega)=\{2,3,4,5,6,7,8,9,10,11,12\}$$

نحسب قانون الاحتمال للمتغير العشوائي Y، نجد أن:

$$IP(Y=2)=IP((1,1))=\frac{1}{36}$$
$$IP(Y=3)=IP((1,2),(2,1))=\frac{2}{36}$$
$$IP(Y=4)=IP((1,3),(2,2),(3,1))=\frac{3}{36}$$
$$IP(Y=5)=IP((1,4),(2,3),(3,2),(4,1))=\frac{4}{36}$$

وهكذا وبنفس طريقة الحساب، نحصل على جدول التوزيع الاحتمالي التالي:

Y	2	3	4	5	6	7	8	9	10	11	12
$IP(Y=y)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

وبالتالي يكون تابع التوزيع للمتغير العشوائي Y هو:

$$F_Y(y)=\begin{cases} 0;\ x<2 \\ IP(x\le 2)=IP(x=2)=\frac{1}{36};\ 2\le x<3 \\ IP(x\le 3)=IP(x=2)+IP(x=3)=\frac{1}{12};\ 3\le x<4 \\ IP(x\le 4)=IP(x=2)+IP(x=3)+IP(x=4)=\frac{1}{6};\ 4\le x<5 \\ IP(x\le 5)=\frac{5}{18};\ 5\le x<6 \\ IP(x\le 6)=\frac{5}{12};\ 6\le x<7 \\ IP(x\le 7)=\frac{7}{12};\ 7\le x<8 \\ IP(x\le 8)=\frac{13}{36};\ 8\le x<9 \\ IP(x\le 9)=\frac{5}{6};\ 9\le x<10 \\ IP(x\le 10)=\frac{11}{12};\ 10\le x<11 \\ IP(x\le 11)=\frac{35}{36};\ 11\le x<12 \\ IP(x\le 12)=1;\ x\ge 12 \end{cases}$$

2.3.3. تابع التوزيع في حالة المتغير العشوائي المستمر

تعريف 2.3.3. (متغير عشوائي ذو كثافة)

نقول أن المتغير العشوائي X ، متغير عشوائي ذو كثافة f ، إذا كان

$$IP(X\in[a,b])=\int_a^b f(x)dx$$

ملاحظة:

$$IP(a<X<b)=IP(a\le X\le b)=IP(a<X\le b)=IP(a\le X<b)$$

تعريف 3.3.3. (تابع الكثافة)

نسمي تابع (دالة) كثافة احتمالية كل تطبيق f معرف بـ:

$$f : IR \to IR^{+}$$
$$x \to f(x)$$

حيث أن

- من أجل كل x من IR ، $f(x) \geq 0$.
- $\int_{-\infty}^{+\infty} f(x)dx = 1$

تعريف 4.3.3. (تابع التوزيع)

إذا كان X متغير عشوائي مستمر فإن F_X يمكن تعريفها كما يلي:

$$F_X(x) = IP(X \leq x) = \int_{-\infty}^{x} f(t)dt \ ; \ \forall x \in IR$$

وفي هذه الحالة يكون قانون احتمال المتغير العشوائي X كما يلي:

$$IP(a \leq x < b) = IP_X([a,b[) = F_X(b) - F_X(a) = \int_{a}^{b} f(t)dt$$

مثال:

ليكن X متغير عشوائي حقيقي، كثافته الاحتمالية هي:

$$f(x) = \begin{cases} 2x & ,0 \leq x \leq 1 \\ 0 & ,\text{ailleurs} \end{cases}$$

نلاحظ أن

- $f(x) \geq 0, \quad \forall x \in IR$
- $x < 0, \quad F_X(x) = \int_{-\infty}^{x} f(u)du = 0$
- $0 \leq x < 1, \quad F_X(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{0} 0du + \int_{0}^{x} 2udu = x^2$

- $x \geq 1, \quad F_X(x) = \int_{-\infty}^{x} f(u)du = \int_{-\infty}^{0} 0du + \int_{0}^{1} 2udu + \int_{1}^{x} 0du = 1$

ومنه يكون تابع التوزيع كما يلي:

$$F_X(x) = \begin{cases} 0 & ; x < 0 \\ x^2 & ; 0 \leq x < 1 \\ 1 & ; x \geq 1 \end{cases}$$

4.3. عزوم المتغير العشوائي

تعريف 1.4.3. (الأمل الرياضي)

- نعرف الأمل الرياضي للمتغير العشوائي X ، أنه حاصل الجمع الناتج عن ضرب كل قيمة للمتغير العشوائي X في الاحتمال المناظر لها $IP(X = x)$. وهذا في حالة X متغير عشوائي متقطع. ونكتب:

$$IE(X) = \sum_i x_i IP(X = x_i)$$

- نعرف الأمل الرياضي للمتغير العشوائي X ، أنه تكامل حاصل ضرب x في دالة الكثافة $f(x)$ ، وهذا في حالة X متغير عشوائي مستمر. ونكتب:

$$IE(X) = \int_{IR} xf(x)dx$$

نظرية 1.4.3. [18]

من أجل كل زوج ثابت a و b ، فإن

$$IE(aX + b) = aIE(X) + b$$

البرهان

- في حالة المتغير العشوائي المتقطع

$$\begin{aligned} IE(aX+b) &= \sum_{x;p(x)>0}(ax+b)p(x) \\ &= a\sum_{x;p(x)>0}xp(x)+b\sum_{x;p(x)>0}p(x) \\ &= aIE(X)+b \end{aligned}$$

- في حالة المتغير العشوائي المستمر

$$\begin{aligned} IE(aX+b) &= \int_{-\infty}^{\infty}(ax+b)f(x)dx \\ &= a\int_{-\infty}^{\infty}xf(x)dx+b\int_{-\infty}^{\infty}f(x)dx \\ &= aIE(X)+b \end{aligned}$$

تعريف 2.4.3. (التباين والانحراف المعياري)

نعرف تباين المتغير العشوائي X بـ:

$$Var(X)=IE\left(\left(X-IE(X)\right)^2\right)$$

نعرف أيضا الانحراف المعياري بـ:

$$\sigma_X=\sqrt{Var(X)}$$

قضية 1.4.3. (خواص التباين) [18]

من أجل كل a و b من IR.

1) $Var(X)\geq 0$

2) $Var(X)=IE(X^2)-\left(IE(X)\right)^2$

3) $Var(aX)=a^2Var(X)$

4) $Var(X+b)=Var(X)$

5) $V(X)=0$ إذا و فقط إذا كان X ثابت $p.s.$

البرهان

1) الخاصية الأولى واضحة.

2) نضع $IE(X)=\mu$ ،

$$\begin{aligned}
Var(X) &= IE\left(\left(X-IE(X)\right)^2\right)\\
&= IE\left(\left(X-\mu\right)^2\right)\\
&= IE\left(\left(X^2-2X\mu+\mu^2\right)\right)\\
&= IE\left(X^2\right)-2IE\left(X\mu\right)+\mu^2\\
&= IE\left(X^2\right)-2IE\left(X\right)\mu+\mu^2\\
&= IE\left(X^2\right)-2\mu^2+\mu^2\\
&= IE\left(X^2\right)-\mu^2\\
&= IE\left(X^2\right)-\left(IE(X)\right)^2
\end{aligned}$$

(3

$$\begin{aligned}
Var(aX) &= IE\left(\left(aX-IE(aX)\right)^2\right)\\
&= IE\left(\left(a\left(X-IE(X)\right)\right)^2\right)\\
&= a^2 IE\left(\left(X-IE(X)\right)^2\right)\\
&= a^2 Var(X)
\end{aligned}$$

(4

$$\begin{aligned}
Var(X+b) &= IE\left(\left(X+b-IE(X+b)\right)^2\right)\\
&= IE\left(\left(X+b-IE(X)+b\right)^2\right)\\
&= IE\left(\left(X-IE(X)\right)^2\right)\\
&= Var(X)
\end{aligned}$$

5) إذا كان $X=c$ ثابت $p.s.$ ، إذن $IE(X)=IE(c)=c$ و $IE(X^2)=IE(c^2)=c^2$ ، ومنه

$$Var(X)=c^2-c^2=0$$

تعريف 3.4.3. (التغاير)

ليكن X و Y متغيرين عشوائيين، حيث أن $Var(X)<+\infty$ و $Var(Y)<+\infty$ ، نعرف تغاير المتغيرين العشوائيين X و Y ، ونكتب $Cov(X,Y)$ بـ:

$$\begin{aligned} Cov(X,Y) &= IE\big((X-IE(X))(Y-IE(Y))\big) \\ &= IE(XY)-IE(X)IE(Y) \end{aligned}$$

ملاحظات:

- $Cov(X,X)=Var(X)$
- إذا كان X و Y مستقلتين، يكون لدينا:

$$Cov(X,Y)=IE(XY)-IE(X)IE(Y)=0$$

قضية 2.4.3. [18]

إذا كان X و Y متغيرين عشوائيين، حيث $IE(X^2)<+\infty$ و $IE(Y^2)<+\infty$. فإن

$$Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$$

البرهان

$$\begin{aligned} Var(X+Y) &= IE\left((X+Y)^2\right)-\left(IE(X+Y)\right)^2 \\ &= IE\left(X^2+2XY+Y^2\right)-\left(IE(X)+IE(Y)\right)^2 \\ &= IE\left(X^2\right)+2IE(XY)+IE\left(Y^2\right)-\left(IE(X)\right)^2-2IE(X)IE(Y)-\left(IE(Y)\right)^2 \\ &= IE\left(X^2\right)-\left(IE(X)\right)^2+IE\left(Y^2\right)-\left(IE(Y)\right)^2+2IE(XY)-2IE(X)IE(Y) \\ &= Var(X)+Var(Y)+2Cov(X,Y) \end{aligned}$$

5.3. التابع المولد للعزوم

تعريف 1.5.3.

نعرف من أجل كل حقيقي t، التابع المولد φ للمتغير العشوائي X بـ:

- في حالة X متغير عشوائي مستمر

$$\varphi_X(t) = IE\left[e^{tX}\right] = \int_{IR} e^{tx} f_X(x)\,dx$$

- في حالة X متغير عشوائي متقطع

$$\varphi(t) = IE\left(e^{tX}\right) = \sum_x e^{tx} p(x)$$

6.3. الدالة المميزة

تعريف 1.6.3.

ليكن X متغير عشوائي حقيقي. نسمي الدالة المميزة لـ X (أو لقانون X) دالة المتغير الحقيقي t المعرفة بـ:

- في حالة X متغير عشوائي مستمر

$$\phi_X(t) = IE\left[e^{itX}\right] = \int_{IR} e^{itx} f_X(x)dx$$

- في حالة X متغير عشوائي متقطع

$$\phi_X(t) = IE\left[e^{itX}\right] = \sum_{k\in I} e^{itx_k} p_k$$

مثال:

اعتمادا علـــــى نتائج تجربة السابقة والمتمثلة في رمي قطعة نقدية ثلاث مرات متتالية، نحسب كل من الأمل الرياضي $IE(X)$ والتباين $Var(X)$ للمتغيــر العشوائي X كما يلي:

لدينا مما سبق:

X	0	1	2	3
IP_X	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

✓ حساب الأمل الرياضي

$$\begin{aligned}E(X) &= \sum_{i=1}^{4} x_i IP(X = x_i)\\ &= x_1 IP(X = x_1) + x_2 IP(X = x_2) + x_3 IP(X = x_3) + x_4 IP(X = x_4)\\ &= (0)\left(\frac{1}{8}\right) + (1)\left(\frac{3}{8}\right) + (2)\left(\frac{3}{8}\right) + (3)\left(\frac{1}{8}\right) = \frac{3}{2}\end{aligned}$$

✓ حساب التباين

لدينا

$$Var(X) = IE(X^2) - IE(X)^2$$

ولدينا

$$\begin{aligned}IE(X^2) &= \sum_{i=1}^{4} x_i^2 IP(X = x_i)\\ &= (0)\left(\frac{1}{8}\right) + (1)\left(\frac{3}{8}\right) + (4)\left(\frac{3}{8}\right) + (9)\left(\frac{1}{8}\right) = 3\end{aligned}$$

وبالتالي

$$Var(X) = 3 - \left(\frac{3}{2}\right)^2 = \frac{3}{4}$$

نحسب كل من الدالة المولدة للعزوم $\varphi(t)$ والدالة المميزة $\phi_x(t)$ كما يلي:

$$\begin{aligned}\varphi(t) &= IE\left(e^{xt}\right) = \sum_{i=1}^{4} e^{x_i t} IP\left(X = x_i\right) \\ &= e^{0t} IP\left(X = 0\right) + e^{t} IP\left(X = 1\right) + e^{2t} IP\left(X = 2\right) + e^{3t} IP\left(X = 3\right) \\ &= \frac{1}{8} + e^{t}\left(\frac{3}{8}\right) + e^{2t}\left(\frac{3}{8}\right) + e^{3t}\left(\frac{1}{8}\right) \\ &= \frac{1}{8}\left(1+e^{3t}\right) + \frac{3}{8}\left(e^{t} + e^{2t}\right)\end{aligned}$$

وبالتالي

$$\phi_X(t) = IE\left(e^{ixt}\right) = \frac{1}{8}\left(1+e^{3it}\right) + \frac{3}{8}\left(e^{it} + e^{2it}\right)$$

نظرية 1.6.3. [18]

ليكن X و Y متغيرين عشوائيين مستقلين. من أجل كل t من IR ، فإن

$$\phi_{X+Y}(t) = \phi_X(t)\phi_Y(t)$$

7.3. التوزيعات الاحتمالية الشهيرة

1.7.3. التوزيعات الاحتمالية للمتغيرات المنفصلة (المتقطعة)

1) توزيع برنولي (Distribution de Bernoulli)

هو أبسط مثال للتوزيع النقطي حيث أن X هنا يأخذ قيمتين فقط هما 0 و 1، أي أن

$$X(\Omega) = \{0,1\}$$

نعتبر A الحدث المتعلق بتجربة معينة واحتماله هو $IP(A) = p$ ، حيث أن:

$$X(A) = 1 \quad \text{و} \quad X(\overline{A}) = 0$$

قانون احتمال X هو

$$\begin{cases} IP(X=1)=IP(A)=p \\ IP(X=0)=IP(\overline{A})=1-p=q \end{cases}$$

ونكتب

$$X \to \mathcal{B}(p)$$

قانون توزيع برنولي يعطى بالشكل

$$f(k)=IP(X=k)=p^k.q^{1-k}=p^k.(1-p)^{1-k},\quad k=0,1$$

قضية 1.7.3. [11]

إذا كان X متغير عشوائي يتبع توزيع برنولي حيث $p\in[0,1]$ ، فإن

$Var(X)=p(1-p)$ و $IE(X)=p$

البرهان

لدينا

$$IE(X)=\sum_{i=1}^{2}x_i p_i=0(1-p)+1(p)=p$$

ومن جهة أخرى، لدينا

$$\begin{aligned} Var(X)&=IE(X^2)-(IE(X))^2 \\ &=(0)^2(1-p)+(1)^2(p)-p^2 \\ &=p-p^2 \\ &=p(1-p) \end{aligned}$$

2) التوزيع الثنائي (Distribution Binomiale)

نعتبر نفس الشروط المتبعة في توزيع برنولي أي أن هناك حدث A متعلق بتجربة معينة واحتماله $IP(A)=p$ ، ونفرض أننا كررنا هذه التجربة n مرة، نعتبر في كل مرة متغير عشوائي X_i حيث $i=1,2,...$.

حيث X_i هو متغير برنولي في المرة i، الأمر الذي يعني أن

$$IP(X_i=0)=IP(\overline{A})=q=(1-p)$$

$$IP(X_i=1)=IP(A)=p$$

من الواضح أن هذه التجارب مستقلة، نعتبر الآن متغير عشوائي جديد هو

$$X=\sum_{i=1}^{n}X_i=X_1+...+X_n$$

حصولنا على k نجاح من هذه التكرارات يقتضي بالضرورة الحصول على $n-k$ فشل وبما أن الحوادث مستقلة فإن حدث حصولنا على k نجاح يتم بعدد من الطرق يساوي C_n^k.

قانون التوزيع الثنائي يعطى كما يلي:

$$f(k)=IP(X=k)=C_n^k p^k.q^{n-k}=C_n^k p^k.(1-p)^{n-k},\quad k=0,1,...,n$$

ونكتب

$$X\to \mathcal{B}(n,p)$$

قضية 2.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون التوزيع الثنائي حيث $p\in[0,1]$ و $n\geq 1$ ، فإن:

$$Var(X)=np(1-p) \text{ و } IE(X)=np$$

3) التوزيع الهندسي (Distribution Géométrique)

نعتبر أن A حدث متعلق بتجربة معينة واحتماله $IP(A)=p$. نستمر في إجراء هذه التجربة حتى يتحقق الحدث A (النجاح).

يكون عندئذ قانون احتمال X هو

$$IP_X(k)=IP(X=k)=q^{k-1}p \qquad ,k\in IN^*$$

ونكتب

$$X\to \mathcal{G}(p)$$

قضية 3.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون التوزيع الهندسي حيث $p\in[0,1]$ ، فإن

$$Var(X)=\frac{1-p}{p^2} \text{ و } IE(X)=\frac{1}{p}$$

4) توزيع بواسون (Distribution de Poisson)

قانون احتمال المتغير العشوائي X هو:

$$IP_X(k)=IP(X=k)=\frac{\lambda^k e^{-\lambda}}{k!} \qquad ,k\in IN,\ \ \lambda>0$$

هذا القانون خاص بالاحتمالات الصغيرة انطلاقا من قانون التوزيع الثنائي وهذا بملاحظة أن تطبيق القانون الثنائي بالأمر الصعب كلما كبر n وصغر المقدار الاحتمالي p ، ويمكن الحصول علية بالتقريب في التوزيع الثنائي كما يلي:

لدينا قانون التوزيع الثنائي معطى بالعبارة التالية

$$IP(X=k)=\binom{n}{k}p^k q^{n-k}, \quad k=0,1,\dots,n$$

بوضع $np=\lambda$ أي أن $p=\frac{\lambda}{n}$

وبالتالي يكون لدينا

$$IP(X=k)=\binom{n}{k}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$=\frac{n(n-1)........(n-k+1)}{k!}\left(\frac{\lambda}{n}\right)^k\left(1-\frac{\lambda}{n}\right)^{n-k}$$

$$=\frac{nn\left(1-\frac{1}{n}\right)n\left(1-\frac{2}{n}\right)....n\left(1-\frac{k-1}{n}\right)}{k!}\frac{\lambda^k}{n^k}\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

$$IP(X=k)=\frac{\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)....\left(1-\frac{k-1}{n}\right)}{k!}\lambda^k\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

$$=\frac{\lambda^k}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)....\left(1-\frac{k-1}{n}\right)\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^k}$$

لما $n\to+\infty$

$$\lim_{n\to+\infty}\left(1-\frac{\lambda}{n}\right)^{n-k}=\lim_{n\to+\infty}\left(1-\frac{\lambda}{n}\right)^n\left(1-\frac{\lambda}{n}\right)^{-k}=e^{-\lambda}.1=e^{-\lambda}$$

ومنه يكون لدينا

$$IP(X=k)=\frac{\lambda^k e^{-\lambda}}{k!},\quad k\in IN$$

ونكتب

$$X\to\mathcal{P}(\lambda)$$

قضية 4.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون توزيع بواسون حيث $\lambda>0$ ، فإن

$$IE(X)=Var(X)=\lambda$$

2.7.3. التوزيعات الاحتمالية للمتغيرات المستمرة

1) التوزيع المنتظم (Distribution Uniforme)

نقول أن المتغير العشوائي X يتبع قانون التوزيع المنتظم على المجال $[a,b]$، إذا كانت كثافته الاحتمالية معطاة كما يلي:

$$f(x)=\begin{cases}\dfrac{1}{b-a} & ,x\in[a,b]\\ 0 & ,x\notin[a,b]\end{cases}$$

ونكتب

$$X \longrightarrow \mathcal{U}([a,b])$$

قضية 5.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون التوزيع المنتظم، فإن:

$$Var(X)=\frac{(b-a)^2}{12} \text{ و } IE(X)=\frac{b+a}{2}$$

2) التوزيع الاسي (Distribution Exponentielle)

نقول أن المتغير العشوائي X يتبع قانون التوزيع الآسي، إذا كانت كثافته الاحتمالية معطاة كما يلي:

$$f(x)=\begin{cases}\lambda e^{-\lambda x} & ,x\geq 0 \qquad (\lambda>0)\\ 0 & ,x<0\end{cases}$$

ونكتب

$$X \longrightarrow \mathcal{E}xp(\lambda)$$

قضية 6.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون التوزيع الأسي، فإن:

$$Var(X)=\frac{1}{\lambda^2} \text{ و } IE(X)=\frac{1}{\lambda}$$

3) التوزيع الطبيعي (Distribution Normale)

نقول أن المتغير العشوائي X يتبع قانون التوزيع الطبيعي، إذا كانت كثافته الاحتمالية معطاة كما يلي:

$$f(x)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad (\sigma>0, \mu\in IR, x\in IR)$$

ونكتب

$$X \longrightarrow \mathcal{N}(\mu,\sigma^2)$$

قضية 7.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون التوزيع الطبيعي، فإن:

$$Var(X)=\sigma^2 \text{ و } IE(X)=\mu$$

4) التوزيع الطبيعي المعياري (Distribution Normale réduite)

التوزيع الطبيعي المعياري هو حالة خاصة من التوزيع الطبيعي، نقول أن المتغير العشوائي Z يتبع قانون التوزيع الطبيعي المعياري إذا كانت كثافته الاحتمالية معطاة كما يلي:

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

حيث $\sigma^2=1$ و $\mu=0$.

يسمى Z المتغير الطبيعي المعياري، ويمكن أن نصل إليه باستعمال التبديل التالي:

$$Z=\frac{X-\mu}{\sigma}$$

ونكتب

$$X \longrightarrow \mathcal{N}(0,1)$$

5) توزيع قاما (Distribution gamma)

نقول أن المتغير العشوائي X يتبع قانون توزيع قاما، إذا كانت كثافته الاحتمالية معطاة كما يلي:

$$f(x)=\begin{cases}\dfrac{\lambda^{p}}{\Gamma(p)}x^{p-1}e^{-\lambda x} & ,x>0 \\ 0 & ,x\leq 0\end{cases} \qquad (\lambda>0, p>0)$$

حيث أن

$$\Gamma(p)=\int_{0}^{+\infty}x^{p-1}e^{-x}dx$$

ونكتب

$$X\longrightarrow\gamma(p,\lambda)$$

قضية 8.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون توزيع قاما، فإن:

$$Var(X)=\frac{p}{\lambda^{2}} \text{ و } IE(X)=\frac{p}{\lambda}$$

6) توزيع كاي-تربيع (Distribution du khi-deux)

نقول أن المتغير العشوائي X يتبع قانون توزيع كاي تربيع (كاي مربع)، إذا كانت كثافته الاحتمالية معطاة كما يلي:

$$f(x)=\begin{cases}\dfrac{1}{2^{\frac{n}{2}}\Gamma\left(\frac{n}{2}\right)}x^{\frac{n}{2}-1}e^{-\frac{x}{2}} & ,x>0 \\ 0 & ,x\leq 0\end{cases}$$

ونكتب

$$X\to\chi_{n}{}^{2}$$

قضية 9.7.3. [11]

إذا كان X متغير عشوائي يتبع قانون توزيع كاي تربيع، فإن

$$Var(X)=2n \text{ و } IE(X)=n$$

مثال تطبيقي شامل

ليكن X متغير عشوائي حقيقي، حيث

$$X \longrightarrow \mathcal{B}(n,p)$$

1) أحسب القيمة المتوسطة $IE(X)$.

2) أحسب التباين $Var(X)=IE(X^2)-(IE(X))^2$.

3) أحسب التابع المولد للعزوم $\varphi_X(t)$.

4) استنتج التابع المميز $\phi_X(t)$.

الحل

لدينا

$$X \longrightarrow \mathcal{B}(n,p)$$

1) حساب القيمة المتوسطة $IE(X)$

$$\begin{aligned} IE(X) &= \sum_{k=0}^{n} kIP(X=k) \\ &= \sum_{k=0}^{n} k\binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=1}^{n} k\binom{n}{k} p^k q^{n-k} \\ &= \sum_{k=1}^{n} \frac{n!}{(k-1)!(n-k)!} p^k q^{n-k} \end{aligned}$$

بوضع $k-1=r$، نجد أن

$$IE(X)=np\sum_{r=0}^{n-1}\frac{(n-1)!}{r!(n-1-r)!}p^{r}q^{n-1-r}=np(p-q)^{n-1}$$

وبالتالي ينتج أن

$$IE(X)=np$$

2) حساب التباين $Var(X)=IE(X^2)-(IE(X))^2$

- ✓ أولا نقوم بحساب $IE\left(X^2\right)$ كما يلي:

$$IE\left(X^2\right)=\sum_{i=0}^{n}k^2C_n^k p^k q^{n-k}$$

نضع $k^2=k(k-1)+k$ ، نجد أن:

$$\begin{aligned}
IE\left(X^2\right)&=\sum_{k=0}^{n}\left(k(k-1)+k\right)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}\\
&=\sum_{k=0}^{n}\left(k(k-1)\right)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}+\sum_{k=0}^{n}kC_n^k p^k(1-p)^{n-k}\\
&=\sum_{k=0}^{n}\left(k(k-1)\right)\frac{n!}{k!(n-k)!}p^k(1-p)^{n-k}+IE(X)\\
&=\sum_{k=2}^{n}\frac{n!}{(k-2)!(n-k)!}p^k(1-p)^{n-k}+IE(X)\\
&=n(n-1)p^2\sum_{k=2}^{n}C_{n-2}^{k-2}p^{k-2}(1-p)^{n-k}+IE(X)\\
&=n(n-1)p^2\left(p+(1-p)\right)^{n-2}+IE(X)\\
&=n(n-1)p^2+IE(X)
\end{aligned}$$

وجدنا سابقا أن:

$$IE(X)=np$$

وبالتالي:

$$\begin{aligned}Var(X) &= n^2p^2 - np^2 + np - (np)^2\\ &= np(1-p)\\ &= npq\end{aligned}$$

ومنه

$$Var(X) = npq$$

3) حساب التابع المولد للعزوم $\varphi_X(t)$

$$\begin{aligned}\varphi_X(t) &= IE\left(e^{tx}\right) = \sum_{k=0}^{n} e^{tk} C_n^k p^k q^{n-k}\\ &= \sum_{k=0}^{n} C_n^k \left(pe^t\right)^k q^{n-k} = \left(pe^t + q\right)^n\end{aligned}$$

4) استنتاج التابع المميز $\phi_X(t)$

$$\phi_X(t) = \varphi_X(it) = \left(pe^{it} + q\right)^n$$

الفصل الرابع

الثنائية العشوائية

من الضروري في العديد من المسائل العملية والنظرية دراسة سلوك عدة متغيرات عشوائية معا في آن واحد. وهذا السلوك لا يتعين بمعرفة سلوك كل المتغيرات العشوائية على انفراد فقط، بل لابد من دراسة سلوكها المشترك.

الأهداف التعليمية للفصل:

يتعرف الطالب في هذا الفصل على:

- ✓ قانون الاحتمال المشترك للثنائية العشوائية
- ✓ قوانين الاحتمال الهامشية
- ✓ قوانين الاحتمال الشرطية
- ✓ استقلال المتغيرات العشوائية
- ✓ التغاير ومعامل الارتباط
- ✓ استبدال المتغير.

4. الثنائية العشوائية

1.4. تعاريف

تعريف 1.1.4.(الثنائية العشوائية)

نسمي ثنائية عشوائية X ، كل تطبيق قابل للقياس من فضاء الاحتمال $(\Omega, \mathcal{F}, IP)$ إلى الفضاء $\left(IR^2, \mathcal{B}_{IR^2}\right)$.

تعريف 2.1.4. (الثنائية العشوائية)

ليكن $(\Omega, \mathcal{F}, IP)$ فضاء احتمال وليكن X_1 و X_2 متغيرين عشوائيين معرفين على الفضاء $(\Omega, \mathcal{F}, IP)$ والتي تأخذ قيما في IR . ليكن X التطبيق المعرف بـ:

$$X : \Omega \to IR^2$$
$$\omega \mapsto X(\omega) = \left(X_1(\omega), X_2(\omega)\right)$$

نقول أن $X = (X_1, X_2)$ هي ثنائية عشوائية.

تعريف 3.1.4. (قانون الاحتمال المشترك)

لتكن $X = (X_1, X_2)$ ثنائية عشوائية معرفة بـ: $X : (\Omega, \mathcal{F}, IP) \to \left(IR^2, \mathcal{F}_{IR^2}, IP_X\right)$

نسمي IP_X قانون احتمال الثنائية $X = (X_1, X_2)$ ، ونكتب:

$$\forall B \in \mathcal{B}_{IR^2}, IP_X(B) = IP(X \in B) = IP(\omega \in \Omega,\ X(\omega) \in B)$$

2.4. الثنائية العشوائية المتقطعة

ليكن X و Y متغيرين عشوائيين متقطعين، حيث

$$X = \{x_1, x_2, x_3, \ldots, x_p\}$$

$$Y = \{y_1, y_2, y_3, \ldots, y_q\}$$

1.2.4. قانون احتمال الثنائية العشوائية المتقطعة

لتكن (X, Y) ثنائية عشوائية متقطعة. قانون احتمال الثنائية العشوائية (X, Y) ، معرف بمجموعة الأعداد p_{ij} $(0 \le p_{ij} \le 1)$ ، حيث

$$p_{ij} = IP\left(X = x_i, Y = y_j\right) = IP\left(\left(X = x_i\right) \cap \left(Y = y_j\right)\right)$$

حيث أن p_{ij} يحقق العلاقة: $\sum_{i=1}^{p}\sum_{j=1}^{q} p_{ij} = 1$

نعرف قانون احتمال الثنائية العشوائية (X,Y) أو قانون الاحتمال المشترك لـ X و Y في شكل جدول كما يلي:

$X \backslash Y$	y_1	...	y_j	...	y_q
x_1	p_{11}	...	p_{1j}	...	p_{1q}
$\vdots$			$\vdots$		$\vdots$
x_i	p_{i1}	...	p_{ij}	...	p_{iq}
$\vdots$			$\vdots$		$\vdots$
x_p	p_{p1}	...	p_{pj}	...	p_{pq}

2.2.4. قوانين الاحتمال الهامشية

إن معرفة قانون الاحتمال المشترك للثنائية العشوائية (X,Y) يسمح بمعرفة قانون الاحتمال الهامشي لـ X و Y (IP_X و IP_Y على الترتيب) كما يلي:

- ✓ قانون الاحتمال الهامشي لـ X

$$\begin{aligned} IP_X &= IP\left(X = x_i\right) \\ &= IP\left(X = x_i, Y = y_1\right) + IP\left(X = x_i, Y = y_2\right) + \ldots + IP\left(X = x_i, Y = y_q\right) \\ &= \sum_{j=1}^{q} p_{ij} = p_{i.} \end{aligned}$$

✓ قانون الاحتمال الهامشي لـ Y

$$\begin{aligned} IP_Y &= IP(Y = x_j) \\ &= IP(X = x_1, Y = y_j) + IP(X = x_2, Y = y_j) + \ldots + IP(X = x_p, Y = y_j) \\ &= \sum_{i=1}^{p} p_{ij} = p_{.j} \end{aligned}$$

حيث أن المقادير $p_{i.}$ و $p_{.j}$ تحقق

$$\sum_{i=1}^{p} p_{i.} = \sum_{j=1}^{q} p_{.j} = 1$$

يعطى قانون الاحتمال الهامشي للمتغيرين X و Y في السطر الأخير والعمود الأخير من الجدول كما يلي:

$X \backslash Y$	y_1	$\ldots$	y_j	$\ldots$	y_q	IP_X
x_1	p_{11}	$\ldots$	p_{1j}	$\ldots$	p_{1q}	$p_{1.}$
$\vdots$			$\vdots$		$\vdots$	
x_i	p_{i1}	$\ldots$	p_{ij}	$\ldots$	p_{iq}	$p_{i.}$
$\vdots$			$\vdots$		$\vdots$	$\vdots$
x_p	p_{p1}	$\ldots$	p_{pj}	$\ldots$	p_{pq}	$p_{p.}$
IP_Y	$p_{.1}$	$\ldots$	$p_{.j}$	$\ldots$	$p_{.q}$	$\sum_{i=1}^{p}\sum_{j=1}^{q} p_{ij} = 1$

مثال:

نرمي قطعة نقود متزنة مرتين متتاليتين و نراهن بدينارين على ظهور الصورة (F) في الرمية الأولى، وإذا حصلنا على (F) نراهن بدينار على ظهور (F) في الرمية الثانية. إذا حصلنا على (P) في الرمية الأولى نراهن بـ 3 دنانير على ظهور (F) في الرمية الثانية.

- نقوم بإيجاد قانون الاحتمال المشترك ومن ثم قوانين الاحتمال الهامشية.
  - ✓ مجموعة الإمكانيات لهذه التجربة هي:

$$\Omega = \{FF, FP, PF, PP\}$$

  - ✓ نرمز بـ X للربح في الرمية الأولى، أي أن:

$$X = \{-2, 2\}$$

  - ✓ نرمز بـ Y للربح في الرمية الثانية، أي أن:

$$Y = \{-1, 1, -3, 3\}$$

$$(X, Y) = \{(2,-1), (2,1), (-2,-3), (-2,3)\}$$

  - ✓ قانون الاحتمال المشترك للثنائية يعطى كما يلي:

$$IP(X = 2, Y = -1) = IP(FP) = \frac{1}{4}$$

$$IP(X = 2, Y = 1) = IP(FF) = \frac{1}{4}$$

$$IP(X = -2, Y = -3) = IP(PP) = \frac{1}{4}$$

$$IP(X = -2, Y = 3) = IP(PF) = \frac{1}{4}$$

نلخص قانون الاحتمال المشترك جدوليا كما يلي:

$X \backslash Y$	1-	1	-3	3
-2	0	0	$\frac{1}{4}$	$\frac{1}{4}$
2	$\frac{1}{4}$	$\frac{1}{4}$	0	0

وبالتالي يكون قانون الاحتمال الهامشي لـ X كما يلي:

$$\begin{aligned} IP(X=-2) &= IP(X=-2,Y=-1)+IP(X=-2,Y=1) \\ &\quad +IP(X=-2,Y=-3)+IP(X=-2,Y=3) \\ &= \frac{1}{4}+\frac{1}{4}=\frac{1}{2} \end{aligned}$$

$$\begin{aligned} IP(X=2) &= IP(X=2,Y=-1)+IP(X=2,Y=1) \\ &\quad +IP(X=2,Y=-3)+IP(X=2,Y=3) \\ &= \frac{1}{4}+\frac{1}{4}=\frac{1}{2} \end{aligned}$$

ويكون قانون الاحتمال الهامشي لـ Y كما يلي:

$$IP(Y=-1)=IP(Y=1)=IP(Y=-3)=IP(Y=3)=\frac{1}{4}$$

نلخص قانون الاحتمال الهامشي لـ X و Y جدوليا كما يلي:

$X \backslash Y$	1-	1	-3	3	IP_X
-2	0	0	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{2}$
2	$\frac{1}{4}$	$\frac{1}{4}$	0	0	$\frac{1}{2}$
IP_Y	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	$\frac{1}{4}$	1

3.2.4. قانون الاحتمال الشرطي للثنائية العشوائية المتقطعة

ليكن X و Y متغيران عشوائيان متقطعان معرفان على فضاء الاحتمال $(\Omega,\mathcal{F},IP)$.

نعتبر الحدثين $A=(X=x_i)$ و $B=(Y=y_j)$ ولدينا

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

حيث $IP(B)\neq 0$

وبالتالي يكون لدينا:

- قانون الاحتمال الشرطي لـ X

$$IP(X=x_i/Y=y_j)=\frac{IP(X=x_i\cap Y=y_j)}{IP(Y=y_j)}=p_{i/j}$$

أي أن $p_{i/j}=\dfrac{p_{ij}}{p_{.j}}$ حيث $p_{.j}\neq 0$.

- قانون الاحتمال الشرطي لـ Y

$$IP(Y=y_j/X=x_i)=\frac{IP(Y=y_j\cap X=x_i)}{IP(X=x_i)}=p_{j/i}$$

أي أن $p_{j/i}=\dfrac{p_{ij}}{p_{i.}}$ حيث $p_{i.}\neq 0$.

مثال:

رمينا قطعة نقود متجانسة ثلاث مرات متتالية. ليكن X عدد مرات ظهور (F) في الرمية الأولى، وليكن Y عدد مرات ظهور (P) في الرميات الثلاث.

لإيجاد قانون الاحتمال الشرطي للثنائية العشوائية (X,Y)، نقوم أولا بإيجاد قانون الاحتمال المشترك للثنائية العشوائية (X,Y) ومن ثم قوانين الاحتمال الهامشية لـ X و Y.

1) قانون الاحتمال المشترك

$$\Omega=\{PPP,PPF,PFP,PFF,FPP,FPF,FFP,FFF\}$$

- X تمثل عدد مرات ظهور (F) في الرمية الأولى، أي أن $X=\{0,1\}$.
- Y تمثل عدد مرات ظهور (P) في الرميات الثلاث، أي أن $Y=\{0,1,2,3\}$.
- إذا اعتبرنا سلوك المتغيرين العشوائيين معا فإن:

$$(X,Y)=\{(0,1),(0,2),(0,3),(1,0),(1,1),(1,2)\}$$

وبالتالي

$$IP(X=0)=\frac{4}{8}=\frac{1}{2},\quad IP(X=1)=\frac{4}{8}=\frac{1}{2}.$$
$$IP(Y=0)=\frac{1}{8},\quad IP(Y=1)=\frac{3}{8},\quad IP(Y=2)=\frac{3}{8},\quad IP(Y=3)=\frac{1}{8}.$$

قانون الاحتمال المشترك لـ X و Y يكون كما يلي:

$$IP(X=0,Y=0)=0,\quad IP(X=0,Y=1)=\frac{1}{8}$$
$$IP(X=0,Y=2)=\frac{2}{8},\quad IIP(X=0,Y=3)=\frac{1}{8}$$

$$IP(X=1,Y=0)=\frac{1}{8},\quad IP(X=1,Y=1)=\frac{2}{8}$$
$$IP(X=1,Y=2)=\frac{1}{8},\quad IP(X=1,Y=3)=0.$$

يمكن تلخيص النتائج السابقة جدوليا كما يلي:

$X \backslash Y$	0	1	2	3	IP_X
0	0	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	$\frac{4}{8}$
1	$\frac{1}{8}$	$\frac{2}{8}$	$\frac{1}{8}$	0	$\frac{4}{8}$
IP_Y	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	1

2) قانون الاحتمال الشرطي للثنائية العشوائية (X,Y)

✓ $p_{i/j}$

$$p_{1/1}=IP(X=x_1/Y=y_1)=IP(X=0/Y=0)=\frac{IP(X=0\cap Y=0)}{IP(Y=0)}=0$$

$$p_{2/1}=IP(X=x_2/Y=y_1)=IP(X=1/Y=0)=\frac{IP(X=1\cap Y=0)}{IP(Y=0)}=\frac{1/8}{1/8}=1$$

$$p_{1/2}=IP(X=x_1/Y=y_2)=IP(X=0/Y=1)=\frac{IP(X=0\cap Y=1)}{IP(Y=1)}=\frac{1/8}{3/8}=\frac{1}{3}$$

$$p_{2/2}=IP(X=x_2/Y=y_2)=IP(X=1/Y=1)=\frac{IP(X=1\cap Y=1)}{IP(Y=1)}=\frac{2/8}{3/8}=\frac{2}{3}$$

$$p_{1/3}=IP(X=x_1/Y=y_3)=IP(X=0/Y=2)=\frac{IP(X=0\cap Y=2)}{IP(Y=2)}=\frac{2/8}{3/8}=\frac{2}{3}$$

$$p_{2/3}=IP(X=x_2/Y=y_3)=IP(X=1/Y=2)=\frac{IP(X=1\cap Y=2)}{IP(Y=2)}=\frac{1/8}{3/8}=\frac{1}{3}$$

$$p_{1/4}=IP(X=x_1/Y=y_4)=IP(X=0/Y=3)=\frac{IP(X=0\cap Y=3)}{IP(Y=3)}=\frac{1/8}{1/8}=1$$

$$p_{2/4}=IP(X=x_2/Y=y_4)=IP(X=1/Y=3)=\frac{IP(X=1\cap Y=3)}{IP(Y=3)}=0$$

ومنه قانون الاحتمال الشرطي $p_{i/j}$ يعطى جدوليا كما يلي:

X \ $p_{i/j}$	$p_{i/1}$	$p_{i/2}$	$p_{i/3}$	$p_{i/4}$
$X=0$	0	$\frac{1}{3}$	$\frac{2}{3}$	1
$X=1$	1	$\frac{2}{3}$	$\frac{1}{3}$	0
$\sum$	1	1	1	1

✓ $p_{j/i}$

$$p_{1/1}=IP(X=x_1/Y=y_1)=IP(Y=0/X=0)=\frac{IP(X=0\cap Y=0)}{IP(X=0)}=0$$

$$p_{2/1}=IP(X=x_1/Y=y_2)=IP(Y=1/X=0)=\frac{IP(X=0\cap Y=1)}{IP(X=0)}=\frac{1/8}{4/8}=\frac{1}{4}$$

$$p_{3/1}=IP(X=x_1/Y=y_3)=IP(Y=2/X=0)=\frac{IP(X=0\cap Y=2)}{IP(X=0)}=\frac{2/8}{4/8}=\frac{2}{4}$$

$$p_{4/1}=IP(X=x_1/Y=y_4)=IP(Y=3/X=0)=\frac{IP(X=0\cap Y=3)}{IP(X=0)}=\frac{1/8}{4/8}=\frac{1}{4}$$

$$p_{1/2}=IP(X=x_2/Y=y_1)=IP(Y=0/X=1)=\frac{IP(X=1\cap Y=0)}{IP(X=1)}=\frac{1/8}{4/8}=\frac{1}{4}$$

$$p_{2/2}=IP(X=x_2/Y=y_2)=IP(Y=1/X=1)=\frac{IP(X=1\cap Y=1)}{IP(X=1)}=\frac{2/8}{4/8}=\frac{2}{4}$$

$$p_{3/2}=IP(X=x_2/Y=y_3)=IP(Y=2/X=1)=\frac{IP(X=1\cap Y=2)}{IP(X=1)}=\frac{1/8}{4/8}=\frac{1}{4}$$

$$p_{4/2}=IP(X=x_2/Y=y_4)=IP(Y=3/X=1)=\frac{IP(X=1\cap Y=3)}{IP(X=1)}=0$$

ومنه قانون الاحتمال الشرطي $p_{j/i}$ يعطى جدوليا كما يلي:

$p_{i/j}$ \ Y	$Y=0$	$Y=1$	$Y=2$	$Y=3$
$p_{j/1}$	0	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$
$p_{j/2}$	$\frac{1}{4}$	$\frac{2}{4}$	$\frac{1}{4}$	0

4.2.4. استقلال متغيرين عشوائيين متقطعين

نقول أن المتغيران العشوائيان X و Y مستقلان إذا وفقط إذا كان:

$$IP\left(X = x_i, Y = y_j\right) = IP\left(X = x_i\right).IP\left(Y = y_j\right)$$

ومن ذلك ينتج أن

$$IP\left(X = x_i / Y = y_j\right) = IP\left(X = x_i\right)$$
$$IP\left(Y = y_j / X = x_i\right) = IP\left(Y = y_j\right)$$

أي أن تحقق الحادثة $\left(X = x_i\right)$ غير مرتبط بتحقق الحادثة $\left(Y = y_j\right)$.

5.2.4. الأمل الشرطي

الأمل الشرطي للمتغير العشوائي المتقطع Y شرط X ، ونكتب $IE\left(Y / X = x\right)$ يعطى كما يلي:

$$IE\left(Y / X = x\right) = \sum_j y_j IP\left(Y = y_j / X = x_i\right)$$

خواص:

1) الخطية: إذا كان a و b عددان حقيقيان

$$IE\left(aY_1 + bY_2 / X = x\right) = aIE\left(Y_1 / X = x\right) + bIE\left(Y_2 / X = x\right)$$

2) نظرية الأمل المطلق

$$IE\left(IE\left(Y / X\right)\right) = IE\left(Y\right)$$

مثال:

ليكن X و Y متغيرين عشوائيين متقطعين مستقلين، يتبعان التوزيع الثنائي $\mathcal{B}(n, p)$. نريد حساب الأمل الشرطي لـ X بشرط $X + Y = m$.

- نقوم أولا بحساب قانون الاحتمال الشرطي $IP\left(X = k / X + Y = m\right)$.

$$\begin{aligned}
IP\left(X=k/X+Y=m\right)&=\frac{IP\left(X=k,X+Y=m\right)}{IP\left(X+Y=m\right)}\\
&=\frac{IP\left(X=k,Y=m-k\right)}{IP\left(X+Y=m\right)}\\
&=\frac{IP\left(X=k\right)IP\left(Y=m-k\right)}{IP\left(X+Y=m\right)}\\
&=\frac{C_n^k p^k\left(1-p\right)^{n-k}C_n^{m-k}p^{m-k}\left(1-p\right)^{n-m+k}}{C_{2n}^m p^m\left(1-p\right)^{2n-m}}\\
&=\frac{C_n^k C_n^{m-k}p^m\left(1-p\right)^{2n-m}}{C_{2n}^m p^m\left(1-p\right)^{2n-m}}\\
&=\frac{C_n^k C_n^{m-k}}{C_{2n}^m}
\end{aligned}$$

وبالتالي

$$IE\left(X/Y=y\right)=\frac{m}{2}$$

6.2.4. التباين الشرطي

التباين الشرطي للمتغير العشوائي Y شرط X، ونكتب $Var(Y/X=x)$، يعطى بالعلاقة التالية:

$$Var(Y/X=x)=IE\left(\left(Y-IE(Y/X=x)\right)^2/X=x\right)$$

3.4. الثنائية العشوائية المستمرة

1.3.4. قانون احتمال الثنائية العشوائية المستمرة

تعريف 1.3.4. (تابع الكثافة)

إذا كان X و Y متغيرين عشوائيين مستمرين فإن $W=\left(X,Y\right)$ هي ثنائية عشوائية مستمرة وتقبل تابع كثافة معرف كما يلي:

$$f:IR^2\longrightarrow IR^+$$

حيث أن

$$\forall (x,y)\in IR^2,\quad f(x,y)\geq 0$$

$$\iint_{IR} f(x,y)dxdy=1$$

يسمى f بتابع الكثافة المشترك لـ X و Y.

مثال:

ليكن التابع f المعرف كما يلي:

$$f(x,y)=\begin{cases} e^{-x-y}, & x\geq 0,\ \ y\geq 0 \\ 0 & ailleurs \end{cases}$$

نبين أن f هو تابع الكثافة المشترك لـ X و Y.

لدينا

$$\forall (x,y)\in IR^2,\quad f(x,y)\geq 0$$

$$\iint_{IR} f(x,y)dxdy=\int_0^{+\infty}\int_0^{+\infty} e^{-x-y}dxdy=\left.e^{-x}\right]_0^{+\infty}.\left.e^{-y}\right]_0^{+\infty}=1$$

وبالتالي ينتج أن f تابع كثافة مشترك لـ X و Y.

تعريف 2.3.4. (تابع التوزيع المشترك)

لتكن (X,Y) ثنائية عشوائية مستمرة وتقبل تابع كثافة f .

تابع التوزيع للثنائية (X,Y) أو تابع التوزيع المشترك لـ X و Y يعطى كما يلي:

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(x,y)dxdy$$

مثال:

نحسب تابع التوزيع المشترك للمثال السابق كما يلي:

$$F(x,y)=\int_{-\infty}^{y}\int_{-\infty}^{x} f(x,y)dxdy$$
$$=\int_{0}^{y}\int_{0}^{x} e^{-x-y}dxdy$$
$$=e^{-x}\Big]_0^x .e^{-y}\Big]_0^y$$
$$=\left(e^{-x}-1\right)\left(e^{-y}-1\right)$$
$$=1-e^{-x}-e^{-y}+e^{-x-y}$$

وبالتالي

$$F(x,y)=\begin{cases}1-e^{-x}-e^{-y}+e^{-x-y}, \; x\geq 0, \;\; y\geq 0\\ 0 \qquad\qquad\qquad\qquad\qquad\qquad ailleurs\end{cases}$$

2.3.4. قوانين الاحتمال الهامشية

لتكن (X,Y) ثنائية عشوائية مستمرة وتقبل تابع كثافة f . تابع التوزيع الهامشي لـ X يستنتج من تابع التوزيع المشترك لـ X و Y كما يلي:

$$F_X(x)=IP(X\leq x)=IP(X\leq x, Y<+\infty)=IP\left(\lim_{y\to+\infty}\{X\leq x, Y\leq y\}\right)$$
$$=\lim_{y\to+\infty} IP(X\leq x, Y\leq y)=\lim_{y\to+\infty} F(x,y)\equiv F(x,+\infty)$$

أي أن:

$$F_X(x)=F(x,+\infty)$$

بنفس الطريقة نحصل على التوزيع الهامشي لـ Y .

$$F_Y(y)=IP(Y\leq y)=IP(X\leq +\infty, Y<y)=IP\left(\lim_{x\to+\infty}\{X\leq x, Y\leq y\}\right)$$
$$=\lim_{x\to+\infty} IP(X\leq x, Y\leq y)=\lim_{x\to+\infty} F(x,y)\equiv F(+\infty,y)$$

أي أن:

$$F_Y(y)=F(+\infty,y)$$

أي أنه بانتهاء أحد المتغيرين X أو Y إلى $+\infty$ فإن التابع $F(x,y)$ يؤول إلى تابع التوزيع للمتغير العشوائي الآخر، ونكتب:

$$F_X(x)=F(x,+\infty)=\int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)dydx$$

$$F_Y(y)=F(+\infty,y)=\int_{-\infty}^{y}\int_{-\infty}^{+\infty} f(x,y)dxdy$$

يكفي إجراء عملية الاشتقاق على $F(x,y)$ للحصول على الكثافة المشتركة $f(x,y)$ كما يلي:

$$f(x,y)=\frac{\partial^2}{\partial x\partial y}F(x,y)$$

ملاحظات:

إذا كانت الثنائية (X,Y) مستمرة، فإن المتغيرين العشوائيين X و Y مستمرين.

باشتقاق تابع التوزيع الهامشي لـ X و Y (F_X و F_Y عل الترتيب) نحصل على تابع الكثافة الهامشي لـ X و Y (f_X و f_Y عل الترتيب) كما يلي:

$$f_X(x)=\frac{\partial}{\partial x}\left(\int_{-\infty}^{x}\int_{-\infty}^{+\infty} f(x,y)dydx\right)=\int_{-\infty}^{+\infty} f(x,y)dy$$

$$f_Y(y)=\frac{\partial}{\partial y}\left(\int_{-\infty}^{y}\int_{-\infty}^{+\infty} f(x,y)dxdy\right)=\int_{-\infty}^{+\infty} f(x,y)dx$$

أي أن الكثافة الهامشية لـ X هي:

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy$$

والكثافة الهامشية لـ Y هي:

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx$$

مثال:

لتكن (X,Y)ثنائية متغيرات عشوائية تتبع قانون احتمال ذو كثافة:

$$f(x,y)=\begin{cases} ax(x+y) & ,0\le x\le 1 \quad ,0\le y\le 2 \\ 0 & ,ailleurs \end{cases}$$

1) إيجاد قيمة a حتى يكون f تابع كثافة

$$\int_0^2\int_0^1 f(x,y)dxdy=1 \Rightarrow \int_0^2\int_0^1 (ax(x+y))dxdy=1 \Rightarrow a=\frac{3}{5}$$

2) حساب $f_X(x)$, $f_Y(y)$

$$f_X(x)=\int_{-\infty}^{+\infty}\frac{3}{5}x(x+y)dy=\frac{3}{5}\int_0^2 x(x+y)dy=\frac{3}{5}\int_0^2 (x^2+xy)dy$$

$$=\frac{3}{5}\left[x^2y+\frac{1}{2}xy^2\right]_0^2=\frac{3}{5}\left[2x^2+\frac{4}{2}x\right]=\frac{6}{5}x(x+1).$$

$$f_Y(y)=\int_{-\infty}^{+\infty}\frac{3}{5}x(x+y)dx=\frac{3}{5}\int_0^1 x(x+y)dx=\frac{3}{5}\int_0^2 (x^2+xy)dx$$

$$=\frac{3}{5}\left[\frac{1}{3}x^3+\frac{1}{2}x^2y\right]_0^1=\frac{3}{5}\left[\frac{1}{3}+\frac{1}{2}y\right]$$

3.3.4. قانون الاحتمال الشرطي للثنائية العشوائية المستمرة

ليكن X و Y متغيرين عشوائيين مستمرين.

- الكثافة الهامشية الشرطية لـ X بشرط Y معرفة كما يلي:

$$f(x/y)=\frac{f(x,y)}{f_Y(y)} \qquad , f_Y(y)\neq 0$$

- الكثافة الهامشية الشرطية لـ Y بشرط X معرفة كما يلي:

$$f(y/x)=\frac{f(x,y)}{f_X(x)} \qquad , f_X(x)\neq 0$$

حيث أن

$$\int_{IR} f(y/x)dy=1 \text{ و } \int_{IR} f(x/y)dx=1$$

مثال:

لتكن (X,Y) ثنائية متغيرات عشوائية مستمرة تتبع قانون احتمال ذو كثافة:

$$f(x,y)=\begin{cases}8xy & ,0\le x\le 1 \quad ,0\le y\le x\\ 0 & ,ailleurs\end{cases}$$

1) حساب الكثافة الهامشية $f_Y(y)$, $f_X(x)$.

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_0^x 8xydy=4xy^2\Big]_0^x=4x^3$$

حيث $0\le x\le 1$، أي أن

$$f_X(x)=\begin{cases}4x^3 & ,0\le x\le 1\\ 0 & ,ailleurs\end{cases}$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_y^1 8xydy=4x^2y\Big]_y^1=4y(1-y^2)$$

حيث $0\le y\le 1$، أي أن

$$f_Y(y)=\begin{cases}4y(1-y^2) & ,0\le y\le 1\\ 0 & ,ailleurs\end{cases}$$

2) حساب الكثافة الشرطية $f(y/x)$, $f(x/y)$

$$f(x/y)=\frac{f(x,y)}{f_Y(y)}=\frac{8xy}{4y(1-y^2)}=\frac{2x}{1-y^2}$$

حيث $0\le y\le x\le 1$، و $y\ne 1$.

$$f(y/x)=\frac{f(x,y)}{f_X(x)}=\frac{8xy}{4x^3}=\frac{2y}{x^2}$$

حيث $0\le y\le x\le 1$، و $x\ne 0$.

4.3.4. استقلال متغيرين عشوائيين مستمرين

لتكن (X,Y) ثنائية عشوائية مستمرة، وليكن $F_{X,Y}(x,y)$ تابع التوزيع المشترك لـ X و Y.

وليكن $F_X(x)$ و $F_Y(y)$ تابعي التوزيع الهامشيين لـ X و Y على الترتيب.

نقول أن المتغيرين العشوائيين X و Y مستقلين إذا كان:

$$\forall(x,y)\in\left(X(\Omega),Y(\Omega)\right);\quad F_{X,Y}(x,y)=F_X(x)\cdot F_Y(y)$$

أي أن

$$\forall(x,y);\quad f(x,y)=f_X(x)\cdot f_Y(y)\Leftrightarrow X \text{ و } Y \text{ مستقلين}$$

مثال:

لتكن (X,Y) ثنائية متغيرات عشوائية مستمرة تتبع قانون احتمال ذو كثافة:

$$f(x,y)=\begin{cases} e^{-(x+y)} & ,x\geq 0 \quad ,y\geq 0 \\ 0 & ,ailleurs \end{cases}$$

نثبت أن X و Y مستقلان.

لدينا

$$f_X(x)=\int_{-\infty}^{+\infty} f(x,y)dy=\int_{0}^{+\infty} e^{-(x+y)}dy=e^{-x}$$

$$f_Y(y)=\int_{-\infty}^{+\infty} f(x,y)dx=\int_{0}^{+\infty} e^{-(x+y)}dx=e^{-y}$$

نلاحظ أن

$$f(x,y)=f_X(x)\cdot f_Y(y)$$

وهذا يعني أن X و Y مستقلان.

5.3.4. الأمل الشرطي

ليكن X و Y متغيرين عشوائيين مستمرين، وليكن $f(x,y)$ تابع الكثافة المشترك للمتغيرين X و Y، ولتكن $f_Y(y)$ دالة الكثافة الهامشية لـ Y حيث أن $f_Y(y)>0$.

نعرف الأمل الشرطي لـ X شرط Y، ونكتب $IE(X/Y=y)$، بـ:

$$IE(X/Y=y)=\int_{-\infty}^{\infty} x f_{X/Y}(x/y)dx$$

مثال:

لتكن الكثافة الهامشية للمتغيرين العشوائيين المستمرين X و Y هي:

$$f(x,y)=\frac{e^{-\frac{x}{y}}\cdot e^{-y}}{y} \qquad ,0<x<\infty,\quad 0<y<\infty$$

نقوم بحساب $IE(X/Y=y)$.

$$IE(X/Y=y)=\int_0^{\infty} x f_{X/Y}(x/y)\,dx=\int_0^{\infty} x\frac{f(x,y)}{f_Y(y)}dx$$

$$=\int_0^{\infty} x\frac{f(x,y)}{\int_0^{\infty} f(x,y)\,dx}dx=\int_0^{\infty} x\frac{\frac{e^{-\frac{x}{y}}\cdot e^{-y}}{y}}{\int_0^{\infty}\frac{e^{-\frac{x}{y}}\cdot e^{-y}}{y}dx}dx$$

$$=\int_0^{\infty} x\frac{\frac{1}{y}e^{-\frac{x}{y}}}{\int_0^{\infty}\frac{1}{y}e^{-\frac{x}{y}}dx}dx=\int_0^{\infty} x\frac{\frac{1}{y}e^{-\frac{x}{y}}}{-e^{-\frac{x}{y}}\Big|_0^{\infty}}dx$$

$$=\int_0^{\infty} x\left(\frac{1}{y}\right)e^{-\frac{x}{y}}dx$$

$$=y$$

4.4. التغاير ومعامل الارتباط

تعريف 1.4.4. (التغاير)

ليكن X و Y متغيرين عشوائيين، نعرف تغاير X و Y، ونكتب $Cov(X,Y)$ بـ:

$$Cov(X,Y)=IE\big((X-IE(X))(Y-IE(Y))\big)$$

نستطيع كتابة التغاير بصيغة أخرى كما يلي:

$$\begin{aligned}Cov(X,Y)&=IE\big((X-IE(X))(Y-IE(Y))\big)\\&=IE\big(XY-XIE(Y)-YIE(X)+IE(X)IE(Y)\big)\\&=IE(XY)-IE(X)IE(Y)-IE(X)IE(Y)+IE(X)IE(Y)\\&=IE(XY)-IE(X)IE(Y)\end{aligned}$$

- إذا كان X و Y مستقلتين، يكون لدينا

$$Cov(X,Y)=IE(XY)-IE(X)IE(Y)=0$$

قضية 1.4.4. [18]

إذا كان X و Y متغيرين عشوائيين، حيث $IE(X^2)<+\infty$ و $IE(Y^2)<+\infty$. فإن

$$Var(X+Y)=Var(X)+Var(Y)+2Cov(X,Y)$$

البرهان

$$\begin{aligned}Var(X+Y)&=IE\big((X+Y)^2\big)-\big(IE(X+Y)\big)^2\\&=IE\big(X^2+2XY+Y^2\big)-\big(IE(X)+IE(Y)\big)^2\\&=IE(X^2)+2IE(XY)+IE(Y^2)-\big(IE(X)\big)^2-2IE(X)IE(Y)-\big(IE(Y)\big)^2\\&=IE(X^2)-\big(IE(X)\big)^2+IE(Y^2)-\big(IE(Y)\big)^2+2IE(XY)-2IE(X)IE(Y)\\&=Var(X)+Var(Y)+2Cov(X,Y)\end{aligned}$$

تعريف 2.4.4.(معامل الارتباط)

ليكن X و Y متغيرين عشوائيين، نعرف معامل الارتباط بين X و Y، ونكتب $\rho(X,Y)$ بـ:

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)}\cdot\sqrt{Var(Y)}}$$

خواص:

1. إذا كان X و Y مستقلين فإن $\rho(X,Y)=0$.
2. $-1\leq\rho(X,Y)\leq 1$.

4.2.1. استبدال المتغيرات (حالة متغيرين عشوائيين)

ليكن X_1 و X_2 متغيرين عشوائيين مستمرين، يتبعان قانون احتمال ذو كثافة مشتركة $f_{X_1,X_2}(x_1,x_2)$.

نضع

$$Y_1=g_1(X_1,X_2) \text{ و } Y_2=g_2(X_1,X_2)$$

حيث أننا نستطيع حل جملة المعادلتين

$$\begin{cases} y_1=g_1(x_1,x_2) \\ y_2=g_2(x_1,x_2) \end{cases}$$

نرمز للحلول بـ:

تابع الكثافة المشترك للمتغيرين العشوائيين Y_1 و Y_2 يعطى كما يلي:

$$f_{Y_1,Y_2}(y_1,y_2)=f_{X_1,X_2}(x_1,x_2)|J|^{-1}$$

$$\begin{cases} x_1=h_1(y_1,y_2) \\ x_2=h_2(y_1,y_2) \end{cases}$$

حيث أن:

$$J\left(x_1,x_2\right)=\begin{vmatrix}\dfrac{\partial g_1}{\partial x_1} & \dfrac{\partial g_1}{\partial x_2}\\ \dfrac{\partial g_2}{\partial x_1} & \dfrac{\partial g_2}{\partial x_2}\end{vmatrix}\equiv\frac{\partial g_1}{\partial x_1}\frac{\partial g_2}{\partial x_2}-\frac{\partial g_1}{\partial x_2}\frac{\partial g_2}{\partial x_1}\neq 0$$

من أجل كل ثنائية $\left(x_1,x_2\right)$.

مثال:

ليكن X_1 و X_2 متغيرين عشوائيين مستمرين، يتبعان قانون احتمال ذو كثافة مشتركة $f_{X_1,X_2}\left(x_1,x_2\right)$

نضع $Y_1=X_1+X_2$ و $Y_2=X_1-X_2$.

نريد إيجاد تابع الكثافة المشترك للمتغيرين العشوائيين Y_1 و Y_2.

نضع $g_1\left(x_1,x_2\right)=x_1+x_2$ و $g_2\left(x_1,x_2\right)=x_1-x_2$.

لدينا

$$J\left(x_1,x_2\right)=\begin{vmatrix}1 & 1\\ 1 & -1\end{vmatrix}\equiv-1-1=-2$$

ولدينا أيضا

$$\begin{cases}y_1=x_1+x_2\\ y_2=x_1-x_2\end{cases}\Rightarrow\begin{cases}x_1=\dfrac{y_1+y_2}{2}\\ x_2=\dfrac{y_1-y_2}{2}\end{cases}$$

وبالتالي نستنتج أن تابع الكثافة المشتركة لـ Y_1 و Y_2 كما يلي:

$$f_{Y_1,Y_2}\left(y_1,y_2\right)=\frac{1}{2}f_{X_1,X_2}\left(\frac{y_1+y_2}{2},\frac{y_1-y_2}{2}\right)$$

مثال تطبيقي شامل:

لتكن الثنائية العشوائية $\left(X,Y\right)$ ذات الكثافة الاحتمالية المعطاة بـ:

$$f\left(x,y\right)=\begin{cases}ke^{-\theta y} & ,0\leq x\leq y<\infty\\ 0 & ailleurs\end{cases}$$

1. أثبت أن $k=\theta^2$.
2. أحسب الكثافة الهامشية لـ X و Y.
3. هل المتغيران العشوائيان X و Y مستقلان؟
4. أحسب $IP(X \le 1, Y \le 1)$ و $IP(X \le 1 / Y \le 1)$.
5. أحسب تابع الكثافة للشعاع العشوائي $(X, Y-X)$.

الحل:

1. إثبات أن $k=\theta^2$

$$\iint_{IR} f(x,y)dxdy = 1 \Rightarrow \iint_{IR} ke^{-\theta y}dxdy = 1$$
$$\Rightarrow k\int_0^{+\infty}\left(\int_x^{+\infty} e^{-\theta y}dy\right)dx = 1 \Rightarrow k\int_0^{+\infty}\left(\frac{1}{\theta}e^{-\theta y}\right)dx = 1$$
$$\Rightarrow k = \theta^2$$

2. حساب الكثافة الهامشية لـ X و Y.

$$f_X(x) = \int_0^{+\infty} f(x,y)dy = \int_x^{+\infty} \theta^2 e^{-\theta y}dy = \theta e^{-\theta x}$$

$$f_Y(y) = \int_0^{+\infty} f(x,y)dx = \int_0^{y} \theta^2 e^{-\theta y}dx = \theta^2 y e^{-\theta y}$$

3. المتغيران العشوائيان X و Y مستقلان؟

$$f_X(x)\times f_Y(y) = \theta e^{-\theta x}\times \theta^2 y e^{-\theta y} = \theta^3 y e^{-\theta(x+y)} \neq f(x,y)$$

ومنه نستنتج أن X و Y غير مستقلان.

4. حساب $IP(X \le 1, Y \le 1)$ و $IP(X \le 1 / Y \le 1)$

$$IP(X \le 1, Y \le 1) = \int_0^1\left(\int_x^1 f(x,y)dy\right)dx = \int_0^1\left(\int_x^1 \theta^2 e^{-\theta y}dy\right)dx$$
$$= \int_0^1 \theta\left(e^{-\theta x} - e^{-\theta}\right)dx = 1-(1+\theta)e^{-\theta}$$

$$IP(X \le 1 / Y \le 1) = \frac{IP(X \le 1 \cap Y \le 1)}{IP(Y \le 1)}$$
$$= \frac{1-(1+\theta)e^{-\theta}}{\int_0^1 f_Y(y)dy} = \frac{1-(1+\theta)e^{-\theta}}{1-(1+\theta)e^{-\theta}} = 1$$

5. حساب تابع الكثافة للشعاع العشوائي $(X, Y-X)$

$$F_Z(z_1, z_2) = IP(X \le z_1, Y - X \le z_2)$$
$$= \int_{-\infty}^{x}\int_{-\infty}^{y} \theta^2 e^{-\theta y} dx dy = \int_0^{z_1}\left(\int_x^{x+z_2} \theta^2 e^{-\theta y} dy\right)dx$$
$$= \int_0^{z_1} \theta e^{-\theta x}\left(1-e^{-\theta z_2}\right)dx = \left(1-e^{-\theta z_1}\right)\left(1-e^{-\theta z_2}\right)$$

ومنه

$$f_Z(z_1, z_2) = \frac{\partial}{\partial z_1 \partial z_2} F_Z(z_1, z_2) = \theta^2 e^{-\theta(z_1+z_2)}$$

المراجع

[1] **السعدي رجال (2008)**، نظرية الاحتمالات مبادئ الحساب الاحتمالي، دروس و تمارين، الجزء الأول، ديوان المطبوعات الجامعية الساحة المركزية - بن عكنون – الجزائر.

[2] **بوعبد الله صالح (2006-2005)**، محاضرات الإحصاء الرياضي، لطلبة كلية العلوم الاقتصادية.

[3] **جبار عبد مضحي (2011)** ، مقدمة في نظرية الاحتمالات، دار الميسرة للنشر والتوزيع-عمان-الطبعة الأولى.

[4] **سلسلة ملخصات شوم (1995)**، (نظريات ومسائل في الاحتمالات) الدكتور سيمور ليبشتز، الدار الدولية للنشر والتوزيع –القاهرة.

[5] **عبد الحفيظ مصطفى (2008)**، نظرية الاحتمالات مبادئ وتطبيقات، الجزء الأول، ديوان المطبوعات الجامعية الساحة المركزية ـ بن عكنون ـ الجزائر.

[6] **عبد الحفيظ مصطفى (2008-11)**، نظرية الاحتمالات مبادئ وتطبيقات، الجزء الثاني، ديوان المطبوعات الجامعية الساحة المركزية ـ بن عكنون ـ الجزائر.

[7] **عبد الحفيظ محمد فوزي مصطفى (1999)**، الاستدلال الإحصائي (1): نظرية التقدير، جامعة ناصر، الجماهيرية العظمى.

[8] **ALI MOUHOUCHE**, lexique de mathématiques, avec les statistiques et l'astronomie, Français- arabe, niveau secondaire et supérieur.

[9] **AVNER BAR-HEN (2002-2003),** COURS de DEUG, probabilités et statistique, Université Aix-Marseille III**.**

[10] **DOMINIQUE FOATA, AIME FUCHS (1998),** Calcul des probabilités cours, exercice et problème corriges.© DUNOD, Paris.

[11] **JEAN-PIERRE LECOUTRE (2003),** statistique et probabilités, 2^{e}édition, DUNOD.

[12] **KACI REDJDAL (1995),** Cour de probabilités, 1 place centrale de- Ben Aknoun- (Alger). Réimpression.

[13] **KHALED KHALDI (2010),** Méthodes statistiques, office des Publications, 11.

[14] **M. METIVIER (1972),** Notions fondamentales de la théorie des probabilités, maitrises de mathématiques, dunod université.

[15] **MICHEL TERRAZA (2014) ,** Inférence statistique et probabilités, De Boeck Supérieur s.a., 1^{re}édition.

[16] **MOUSSEDEK BOUSSEBOUA (2006),** Eléments de la théorie des probabilités –Tome 2 - office des publications Universitaires, 1 place centrale de – Ben Aknoun –(Alger) ,2 Edition.

[17] **SAADI N (2018-2019),** Polycopié de cours: Variables aléatoires à plusieurs dimensions, UNIVERSITE A.MIRA DE BEJAIA

[18] **SHELDON M. ROSS (1987),** Initiation aux probabilités, Presses polytechniques romandes, 1^{re}édition.