علم الإحصاء:

❖ تعاريف ومصطلحات:

1- علم الإحصاء:

وهو أحد أهم فروع الرياضيات وهو من أهم الفروع وهو يستخدم في كل مجالات الحياة كما ويدخل علم الاحصاء في كل العلوم الأخرى وكل العلوم مهما كانت تحتاج إلى علم الإحصاء **فمثلاً:**(العلوم العلمية – العلوم الإنسانية والأدبية – العلوم الرياضية – الفنون – العلوم العسكرية – العلوم الاجتماعية) لذلك يعتبر علم الإحصاء هو الأساس في كل العلوم.

- **ما هي وظيفة علم الإحصاء(بماذا يهتم علم الإحصاء):**

يهتم علم الإحصاء بشكل أساسي في جمع وتنظيم وتلخيص وعرض وتحليل وتفسير ونشر البيانات الإحصائية.

2- البيانات Data:

هي جزء من المعلومات الموثوق والمدعم بالأرقام أو معلومات تكون مرتبة ومنظمة ومخزنة بطريقة سهلة التداول و الفهم والتفسير ومدعمة بالأرقام ويمكن تحليلها وتفسيرها وليست كل المعلومات بيانات.

لماذا يعتبر علم الإحصاء ضروري لكل العلوم؟:

كل علم من العلوم فيه بيانات ويتعامل مع البيانات وبالتالي نحتاج إلى الإحصاء ليقوم بمعالجة وتحليل وتقييم وتفسير البيانات .

أو بتعبير أخر كل العلوم أساسها البيانات وهي تحتاج إلى دراسات وأبحاث تطبيقية واختبارات وكل بحث أو اختبار أو تجربة يحتاج إلى علم الإحصاء.

من أين نحصل على البيانات:؟

1– تأتي البيانات بشكل أساسي من مراقبة الظواهر ومشاهدة الظواهر الموجودة بالطبيعة.

2– من خلال إجراء التجارب والاختبارات.

3–من خلال الأبحاث والدراسات والمقابلات.

فوائد علم الإحصاء:

1–يمكن تحليل ومعالجة البيانات.

2–تفسير النتائج للأبحاث والتجارب والاختبارات.

3– الكشف عن صلاحية وجودة البيانات لأي بحث ولأي منتج ولأي تجربة جديدة.

4–تقييم أو تقويم الأبحاث والدراسات.

3- المجتمع Population:

عدد لانهائي من الإفراد أو العناصر التي تتعايش مع بعضها البعض وتتميز بخصائص ومواصفات تميزها عن بقية المجتمعات حيث أنه لكل مجتمع خصائص ومميزات تميزه عن المجتمعات الأخرى.

المجتمع البشري يقسم إلى مجتمعات صغيرة (إحصائية) هذا المجتمع الإحصائي معروف عدده وتميزه مجموعة من الصفات الزمانية و المكانية وهذا النوع هو الذي يستخدم في دراسات الأبحاث.

أهم ميزة في المجتمع الإحصائي هو عدد الأفراد (**n**) والمتوسط (**X**).

4- العينة Sample:

هي جزء من المجتمع يجب ألا يقل عدد أفرادها عن (2 – 10 %) من عدد أفراد المجتمع وهي تؤخذ بطريقة عشوائية **وتوجد منها الأنواع التالية:**

أ- العينة العشوائية البسيطة:

هي أصغر وأبسط أنواع العينات تؤخذ بطريقة عشوائية ويجب أن لا يقل عدد أفرادها عن (30) فرد أو عنصر.

ب- العينة العشوائية الطبقية:

وهي تؤخذ من المجتمع إذا كان على شكل طبقات أو إذا كان الهدف هو دراسة طبقات المجتمع.

مثال: المجتمع العربي طبقات من ناحية الغنى (طبقة الأثرياء – طبقة الأغنياء – طبقة الوسط – طبقة الفقراء – طبقة المعدمين).

عند دراسة المجتمع : من كل طبقة نأخذ عينة حسب حجم الطبقة .

ج- العينة العشوائية المنتظمة:

تؤخذ بطريقة عشوائية ولكن في فترات منتظمة أو من أماكن منتظمة.

مثال: إذا أردنا أخذ عينات من مجتمع معين مثلاً شركة تنتج معجون أسنان وأردنا اختبار المنتجات فنقوم مثلا بأخذ العينات كل ساعة أو كل يوم أو كل أسبوع .

مثال:التنظيم في الجامعات (الجامعة – الكلية – السنوات الدراسية –الشعب– الفئات – المجموعات).

5- المتغيرات (المتحولات) V arıable:

هي توابع ذات قيم متغيرة تتعلق دائماً بصفة معينة **وهي على نوعين:**

أ–المتغيرات الكمية. **– ب–**المتغيرات النوعية.

أ– المتغيرات الكمية:

هي المتغيرات التي قياسها بالقياسات المعروفة مثل الطول والوزن والحجم ويمكن التحكم بها.

ب– المتغيرات النوعية:

هي المتغيرات التي لا يمكن قياسها بالطرق المعروفة وهي عبارة عن تدريجات أو مستويات مثل اللون أو درجة الذكاء.

كما يمكن تقسيم المتغيرات من حيث الاحتمال إلى نوعين :

أ– المتغيرات العشوائية:

هي المتغيرات التي لا يمكن التحكم بها (مثل: درجة الحرارة – نسبة الرطوبة – سرعة الرياح).

ب– المتغيرات اللاعشوائية (النظامية):

يمكن التحكم بها وتحديدها مسبقاً(مثل كمية السماد المضافة ، جرعات من الدواء المأخوذة،عدد الريات).

6- الصفة المدروسة category:

هي الصفة التي يختارها الباحث ليدرس تغيرات هذه الصفة أثناء التجربة أو الاختبار.

مثال:من الصفات المدروسة (صفة الإنتاج أو الإنتاجية – طول النبات – لون الثمرة – حجم الثمرة).

7- القيمة الإحصائية Value:

هي القياس أو الرقم الذي نحصل عليه عند قياس الصفة المدروسة ومجموع القيم الإحصائية نسميها البيانات (Data).

8- التوزيع Distribution:

هو طريقة توزع القيم الإحصائية أو طريقة توزع البيانات للمجتمع أو العينة واحتمالاتها.

- **لدينا ثلاث أنواع من التوزيعات:**

1– التوزيع الثنائي:هو التوزيع الذي يمثل قيم من نوعين فقط أو احتمالين فقط مثل (صح،خطأ).

2– التوزيع الطبيعي:هذا التوزيع يمثل توزيع قيم المجتمعات الطبيعية مثل المجتمع البشري.

3– توزيع بواسون: يمثل هذا التوزيع توزيع الحوادث النادرة (الصقيع -الزلازل – البراكين).

يوجد للتحليل الإحصائي **نوعان:**

1- التحليل الإحصائي الأولي Primary:

نحتاج دائماً إلى إجراء تحليل إحصائي أولي لكل البيانات التجريبية أو بيانات بحث معين.

ماذا يشمل التحليل الإحصائي الأولي:

1-وصف البيانات بشكل نظري.

2- حساب المؤشرات الإحصائية الأساسية **Descriptive Statistics**

وتتضمن المؤشرات التالية:

- **أ**-المتوسط أو المعدل الحسابي **Main** أو **Average**.
- **ب**-الوسيط الحسابي **Median**.
- **ج**-المنوال أو القمة **Mode**.
- **د**-الخطأ القياسي Standard Error.
- **هـ**-التباين أو التشتت **Variation**.
- **و**-معامل الاختلاف Coefficient of Variation.

3- إنشاء جدول التوزيع التكراري.

4-العرض البياني للبيانات.

❖ مقاييس النزعة المركزية:

يوجد مقاييس خاصة لقياس النزعة المركزية وهي:

1- المتوسط أو المعدل الحسابي Avarage أو Main.

2- الوسيط Median.

3- المنوال أو القمة Mode.

عامل الاختلاف COFFICIENT OF VARIATION : يرمز عادة لعامل الاختلاف بـ C.V وهو يعرف بأنه أحد مقاييس التشتت أو التباين الهامة والذي يستخدم بشكل عام للمقارنة بين عينتين أو أكثر عندما لا تكون العينات مأخوذة من مجتمع واحد بل من مجتمعات مختلفة ، وهو أي عامل الاختلاف يعطى بالعلاقة التالية :

$$C.V = (S / \grave{X}) * 100$$

حيث أن :

C.V : عامل الاختلاف

S : الانحراف المعياري

$\grave{X}$: متوسط العينة

◄من الجدير بالذكر انه كلما كانت قيمة عامل الاختلاف C.V أصغر كلما دل ذلك على ثبات العينة وتجانس أفرادها وعناصرها أقل تبعثرا مقارنة بالعينة الأخرى الأكبر بقيمة عامل الاختلاف

مثــال ( ) : قارن بين العينتين التاليتين علما أنهما من مجتمعين مختلفين :

X	3000	4500	4000	4100
Y	900	1200	1700	2200

الحــل : للمقارنة بين العينتين باعتبار أن العينات ليست مأخوذة من مجتمع واحد بل من مجتمعات مختلفة ، فان عامل الاختلاف يعطى بالعلاقة التالية :

$$C.V = (S / \grave{X}) * 100$$

وبتطبيق العلاقة السابقة الذكر نصل على قيمة عامل الاختلاف وهي :

$$C.V (X) = 637.7/3900 * 100$$

$$= 16 \%$$

$$C.V (Y) = 439.7/1500 * 100$$

$$= 29 \%$$

ومن الواضح أن عامل الاختلاف للعينة الاولى أصغر من عامل الاختلاف للعينة الثانية وبالتالي فان العينة الاولى أكثر تجانسا من العينة الثانية.

◄ من الجدير بالذكر انه بالرغم من أن الانحراف المعياري للعينة الثانية أصغر من الانحراف المعياري للعينة الاولى فإننا لا نقدر أن نحكم على مدى تجانس العين لان العينات ليست مأخوذة من مجتمع واحد بل من مجتمعات مختلفة ، وهذا ما يؤكد كل ما ذكرناه سابقاً.

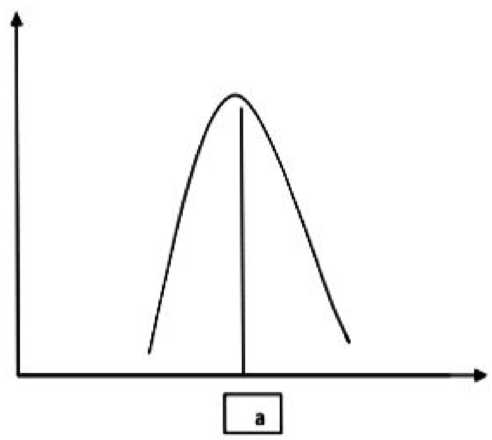
5-6- التوزيع الطبيعي :

5-6-1- تعريف : نقول عن متحول عشوائي x أن توزيعه الاحتمالي يتبع التوزيع الطبيعي إذا كان تابع التوزيع الاحتمالي له على الشكل التالي :

$$f(x) = (1 * e^{-(x-a)/2\delta^2}) / \delta * \sqrt{2\pi}$$

يعتبر التوزيع الطبيعي من التوزيعات الرياضية المعروفة والمهمة من الناحية التطبيقية .

5-6-2- منحنى التوزيع الطبيعي :يملك منحنى التوزيع الطبيعي شكلا ناقوسيا يشبه شكل الجرس على الشكل التالي :



ويمثل هذا الشكل السلوك الطبيعي للظواهر الطبيعية .حيث تبدأ من نقطة منخفضة ثم تزداد حتى تصل إلى الذروة ثم تعود وتنخفض إلى نفس المستوى السابق ،وهذا ما يطلق عليه قانون تناقص الغلة في الاقتصاد(كما ورد في كتاب الاقتصاد الزراعي في السنة الاولى).

ويحسب التوقع الرياضي للتوزيع الطبيعي من العلاقة التالية :

$$M(x) = a$$

مثـــــال (2) : تم رش 1000 شجرة في أحد البساتين لمعالجة أحد الأمراض علما أن فعالية المبيد النظرية هي 98% .والمطلوب هو حساب ما يلي :

1. ما هو احتمال أن تكون ثلاث أشجار على الأقل مصابة؟
2. ما هو احتمال أن يوجد شجرتين مصابتين على الأكثر؟
3. ما هو ؟

الحــــــل : لدينا مسبقا الافتراضات التالية:

$$P = 1- q$$

$$n = 1000$$

$$P(x=r) = (r^n) * P^r * q^{n-r}$$

$$\binom{n}{r} = n!/(n-r!)*r!$$

$$a = n * p = 20$$

$$\delta = \sqrt{n*p*q} = 4.43$$

الآن نطبق العلاقات المذكورة مع مراعاة القيم المسبقة لكل متغير :

1. احتمال أن تكون ثلاث أشجار على الأقل مصابة:

$$P(x>=3) = 0.5 + p(3<x<20)$$

$$p(3<x<20) = \phi(20-20/4.43)- \phi(3-20/4.43)$$

$$p(3<x<20) = 0.4999$$

$$P(x>=3) = 0.5 +4.999 = 0.9999$$

2. احتمال أن يوجد شجرتين مصابتين على الأكثر:

مثـــــال (2) : تم رش 1000 شجرة في أحد البساتين لمعالجة أحد الأمراض علما أن فعالية المبيد النظرية هي 98% .والمطلوب هو حساب ما يلي :

1. ما هو احتمال أن تكون ثلاث أشجار على الأقل مصابة؟
2. ما هو احتمال أن يوجد شجرتين مصابتين على الأكثر؟
3. ما هو ؟

الحـــــل : لدينا مسبقا الافتراضات التالية:

$$P = 1- q$$

$$n = 1000$$

$$P(x=r) = (r^n) * P^r * q^{n-r}$$

$$(\overset{n}{r}) = n\,!/(n-r!)*r!$$

$$a = n * p = 20$$

$$\delta = \sqrt{n*p*q} = 4.43$$

الآن نطبق العلاقات المذكورة مع مراعاة القيم المسبقة لكل متغير :

1. احتمال أن تكون ثلاث أشجار على الأقل مصابة:

$$P(x>=3) = 0.5 + p(3<x<20)$$

$$p(3<x<20) = \phi(20-20/4.43)- \phi(3-20/4.43)$$

$$p(3<x<20) = 0.4999$$

$$P(x>=3) = 0.5 +4.999 = 0.9999$$

2. احتمال أن يوجد شجرتين مصابتين على الأكثر:

$$P(x<=2) = 0.5 – p(2<x<20)$$

$$p(2<x<20) = \phi(20-20/4.43)- \phi(2-20/4.43)$$

$$p(2<x<20) = 0.49997$$

أ- في حالة كون الحوادث متنافية

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

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

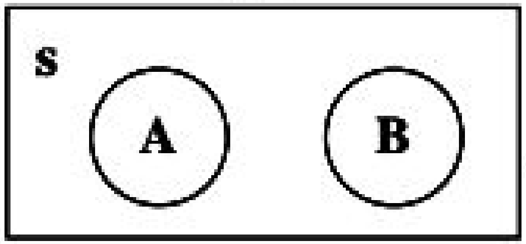
فإذا كان B ، A حادثين متنافيين كما في الشكل (١)

فإن $P\ (A$ أو $B) = P(A) + P(B)$

يرمز أيضاً لاحتمال وقوع أحد الحادثين بالرمز

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

الشكل (١)



حوادث متنافية

ومما سبق نجد أن

إذا كان A , B حادثين شاملين ومتنافين فإن

$$P(A) + P(B) = 1$$

إذا كان A , B حادثين شاملين ومتنافين ومتماثلين فإن

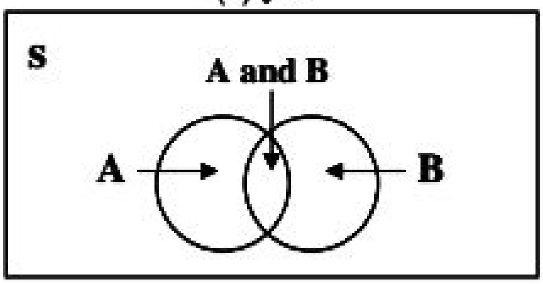
$$P(A) = P(B) = 1/2$$

إذا رمزنا للحادث عدم وقوع A بالرمز $\overline{A}$ فإن الحادثين A و $\overline{A}$ يكونان متنافيين وشاملين ، مثل أن يكون الشخص مدخنا أو غير مدخن

ب - <u>في حالة كون الحوادث غير متنافية</u>

عند عدم اشتراط تنافي الحادثين A و B يكون المقصود بالحادث (A أو B) وقوع A على انفراد أو وقوع B على انفراد أو وقوع الحادثين A و B معا في وقت واحد كما يتضح من الشكل (٢) التالي

شكل (٢)



حوادث غير متنافية

الآن $P(A) + P(B)$ تمثل مجموع الحالات المواتية للحادث A مضافاً إليها مجموع الحالات المواتية للحادث B ولكن يجب ملاحظة أن كل من الحالات المواتية للحادث A وتلك المواتية للحادث B تتضمن الحالات المواتية لوقوع A و B معاً ، وبهذا فإنه في حالة جمع $P(A)$ و $P(B)$ فإننا نجمع

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

إذا كان لدينا الحادثين المستقلين A و B فإن

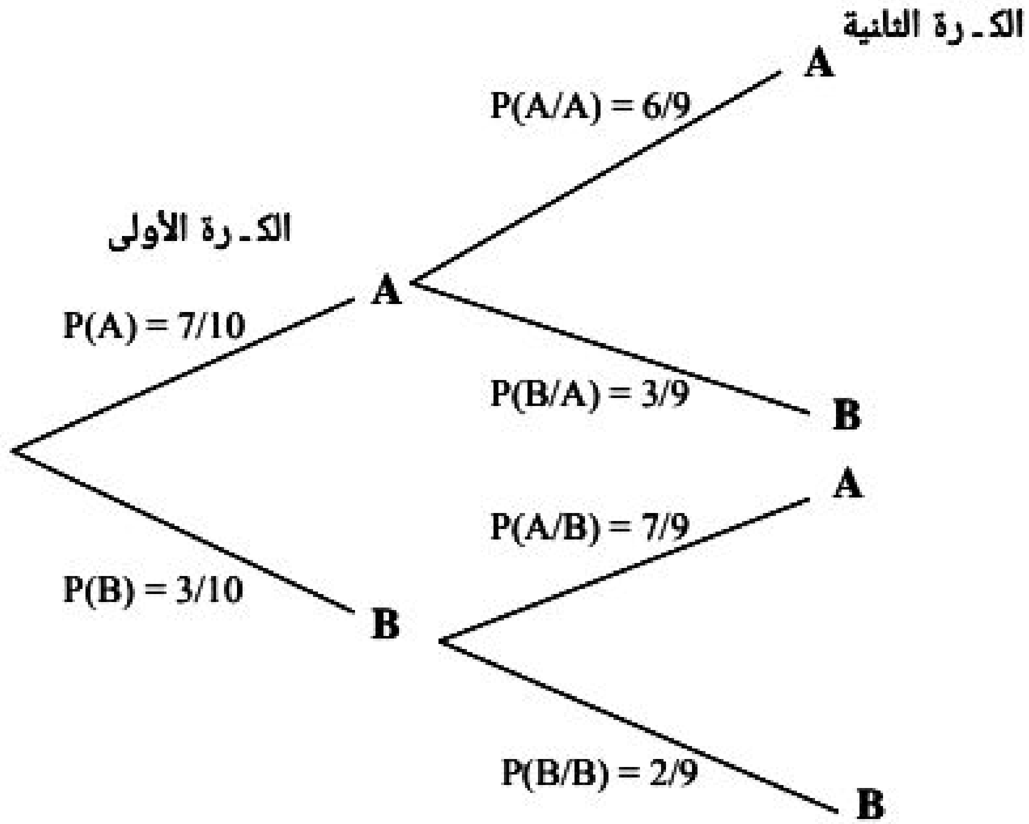
$$P(A \cap B) = P(A)\,P(B)$$

كيس يحتوي على ٣ كرات سوداء و ٧ كرات بيضاء

لنفرض أننا سحبنا منه كرتين كل على حده وبدون إعادة . فإذا رمزنا إلى سحب كرة بيضاء بالرمز A وإلى سحب كرة سوداء بالرمز B فإن الحالات المركبة التي نحصل عليها تتمثل بما يلي:

- الكرتان بيضاوان AA
- الأولى بيضاء والثانية بيضاء AB
- الأولى سوداء والثانية بيضاء BA
- الكرتان سوداوان BB

احتمالات النتائج لسحب الكرة الأولى ثم الثانية بالترتيب يمكن إيضاحها بالرسم التالي:



وهكذا فإن الاحتمال المركب يمكن حسابه كما يلي:

احتمال ان تكون الكرة الأولى بيضاء والثانية بيضاء = احتمال أن تكون الأولى بيضاء مضروبا في احتمال أن تكون الثانية بيضاء بشرط أن تكون الأولى بيضاء.

$$P(A \cap A) = P(A, A) = P(A)\ P(A/A)$$
$$= \frac{7}{10} \times \frac{6}{9} = \frac{42}{90}$$

وبالمثـل

$$P(A \cap B) = P(A)\ P(B/A) = \frac{7}{10} \cdot \frac{3}{9} = \frac{21}{90}$$

$$P(B \cap A) = P(B)\ P(A/B) = \frac{3}{10} \cdot \frac{7}{9} = \frac{21}{90}$$

$$P(B \cap B) = P(B)\ P(B/B) = \frac{3}{10} \cdot \frac{2}{9} = \frac{6}{90}$$

والآن نعمم هذا المثال ونقدم التعريف التالي للاحتمال الشرطي:

- <u>الاحتـ ـمال الشـ ـرطي</u>

إذا كان لدينا الحادثين B , A وكان P (B) لا يساوي الصفر فأن الاحتمال الشرطي للحادث A بشرط وقوع الحادث B يعطي بالمعادلة التالية:

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

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

- **نظريـ ـ ـ ـة بيـ ـ ز (Bayes' Theorem)**

لو افترضنا أن لدينا صندوقين بداخلها مفردات سليمة ومعيبة. مفرده معيبه سحبت عشوائياً من أحد الصندوقين دون تعيين، ونحب أن نعرف ما هو احتمال أن تكون هذه المفردة المعيبة قد سحبت من الصندوق الأول.

للإجابة على أسئلة من هذا النوع نستخدم قاعدة أو نظرية بيز والتي يمكن اعتبارها تطبيقاً للاحتمال الشرطي.

تهدف نظرية بيز إلى حساب احتمالات صحة الفروض بناء على معلومات ميدانية أو تجريبية. في الغالب يكون لدى رجل الأعمال معلومات إضافية عن حادث معين أو مسالة معينة ، إما من خلال خبرته الشخصية أو من خلال ماضي هذا الحادث أو المسألة. الاحتمالات التي تعتمد على الخبرة الشخصية وقبل الحصول على نتائج التجربة تدعى بالاحتمال القبلي Prior Probability فمثلاً الاحتمالات المعتمدة على سجلات المبيعات السابقة أو على الرقم السابق لكمية الإنتاج غير السليم تعتبر أمثلة للاحتمال القبلي. أما عندما يحسب الاحتمال على ضوء معلومات ميدانية مكتسبة يسمى بالاحتمال البعدي Posterior Probability .

فمثلا عند دراسة حجم الأسرة الشائع قد يكون في ذهننا عدة فروض خاصة بهذا الحجم ، هذه الفروض هي:

A1	أن يكون حجم الأسرة الشائع	١
A2	أن يكون حجم الأسرة الشائع	٢
A3	أن يكون حجم الأسرة الشائع	٣
..	.. ..	-
..	.. ..	-
..	.. ..	-
A10	.. ..	10

وهذه الحالات تمثل مجموعة شاملة ومتنافية.

إذا جمعت بيانات ميدانية B عن هذه الظاهرة واستطعنا أن نحسب P(A4/B) فإن هذا الاحتمال يسمى بعديا. مع أن قاعدة بيز يمكن استخدامها لأكثر من حادثين شاملين ومتنافيين إلا أننا ومن أجل التبسيط سنتعرض لتطبيق النظرية على حادثين شاملين متنافيين فقط.

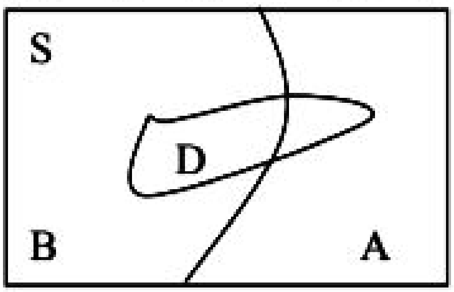
- نظريـة بـيـز

إذا كان A و B حادثين شاملين ومتنافيين في الفراغ العيني S

و D أي حادث في نفس الفراغ بحيث $P(D) \neq O$ فان

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

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



المتغيرات العشوائية والتوزيعات الاحتمالية

RANDOM VARIABLES AND PROBABILITY DISTRIBUTIONS

مر معنا في الفصل السابق عدة تجارب مثل حذف زهرة النرد أو رمي قطعة النقد أو سحب كره أو سحب ورقة من أوراق اللعب ... الخ ، ان النتائج التي يمكن الحصول عليها بتجارب عشوائية، رقمية كانت كما في حالة النرد أو نوعية كما في الكرات وقطعة العملة وورق اللعب وغيرها تدعى بالمتغيرات العشوائية.

- المتغير العشوائي **هو المتغير الذي يتم الحصول على قيمته نتيجة لتجربه عشوائية**

المتغير العشوائي يمكن أن يكون متقطعا DISCRETE أو متصلا CONTINOUNS

المتغير المتقطع:

هو المتغير العشوائي الذي يأخذ قيمة من قيم الأعداد الصحيحة مثل ١ ، ٢ ، ٣ ، ...

أمثلة على المتغير المتقطع :

- عدد أفراد الأسرة في عينة إحصائية مكونه من عده أسر
- عدد الأهداف المسجلة لفريق رياضي خلال الدورى العام
- عدد السيارات المباعه في الشهر لإحدى شركات السيارات ... الخ

المتغير المتصل:

هو المتغير العشوائي الذي يأخذ أي قيمة داخل مدى معين مثل مقاييس الطول والوزن والزمن والقيمة.

أمثلة على المتغير المتصل:

- المبيعات الشهرية من الحليب لإحدى مؤسسات الألبان .
- أطوال طلبة الجامعة .
- أوزان طلاب الصف الأول الابتدائي في إحدى المدارس.
- أثمان البرميل الواحد للنفط الكويتي خلال السنة الأخيرة ... الخ.

- **التوزيـ ـ ـعات الاحتماليـ ـة المتقطعـة Discrete Probability Distributions**

إذا رمينا زهره نرد فإنه يمكننا ان نمثل النتائج الممكنة واحتمالات كل منها بالجدول التالي:

الوجه	١	٢	٣	٤	٥	٦
الاحتمال	١/٦	١/٦	١/٦	١/٦	١/٦	1/6

وإذا رمينا زهرتي نرد وكنا مهتمين بمجموع وجهي الزهر فانه يمكننا أن نمثل النتائج الممكنة واحتمال كل منها بالجدول التالي:

المتغير العشوائي مجموع الوجهين	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
الاحتمال	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

هذا الجدول الذي يتضمن قيم المتغير الشعوائي والاحتمالات المناظرة لها يسمى بالتوزيع الاحتمالي للمتغير العشوائي المتقطع ، وهكذا فإن :

التوزيع الاحتمالي لمجموعة القيم المتقطعة التي يأخذها متغير عشوائي متقطع هو الجدول أو القائمة التي تتضـمن جميـع القيم الممكنة لهذا المتغير مع الاحتمال الخاص بكل منها.

- دالة كثافة الاحتمال **Probabily Density Function (P.D.F.)**

إذا كان لدينا متغير متقطع X يأخذ قيما مختلفة x_1 , x_2 , x_3 ... x_n باحتمال $P(X=x_i)$ حيث $i = 1, 2 ...n$ فإن هذا الاحتمال يسمى بالداله الاحتمالية أو كثافة الاحتمال ويرمز لها بالرمز $f(x_i)$.

وهكذا فإن دالة كثافة الاحتمال للمتغير العشوائي x هي الاحتمال بأن يأخذ x أي قيمة من قيمة الممكنة أي $P(X=xi)$ ويرمز لها بالرمز $f(x)$ على أن يتحقق ما يلي: i) $f(x) \geq 0$ ii) $\sum_{i=1}^{n} f(x_i) = 1$

وتأخذ داله كثافة الاحتمال للمتغير المتقطع X الشكل التالي:

X = x	x_1 x_2 ....x_n

- **دالـة التوزيع الاحتمالي Probability Distribution**

لكننا في الغالب لا نهتم فقط باحتمال أخذ المتغير العشوائي لقيمة واحدة من قيمة الممكنة وإنما نهتم أيضاً باحتمال أخذ المتغير X لقيمة أقل من أو تساوى قيمة ما من قيمة الممكنة وهذه الدالة ندعوها بداله التوزيع الاحتمالي للمتغير X .

دالة التوزيع الاحتمالي للمتغير العشوائي **X** هي الاحتمال بأن **X** أقل من

أو تساوي **x** أي $P(X \le x)$ ويرمز لها بالرمز **F(x)** ويكون

$$F(x) = P(X \le x) = \sum_{xi \le x} f(xi)$$

- **التوزيعات الاحتمالية المتصلة Continuous Probability Distributions**

يمكننا أيضاً أن نعرف التوزيع الاحتمالي لمتغير عشوائي متصل بأنه الجدول أو القائمة الذي يبين لنا القيم المختلفة لهذا المتغير المتصل والاحتمال الخاص بكل قيمة من هذه القيم.

<u>مثـال ٢ :</u>

التوقـــــع ، التبايـــن ، والتغـايـــر

يمكن عن طريق بعض المعلمات (البارامترات) أن نحدد صفات المجتمع الإحصائي وأن نقارن المجتمعات الاحصائيه بعضها ببعض ، ومن أهم هذه المعلمات : التوقع وهو يحدد مركز الموضع للمجتمع والتباين وهو مقياس الانتشار والتغاير وهو مقياس للارتباط بين متغيرين .

وفيما يلي نتناول دراسة هذه المعلمات :

- ### التوقـع Expectation

إذا كان x متغير عشوائي كثافة احتماله f(x) فيعرف المقدار

$$\int_{-\infty}^{\infty} x\, f(x)\, dx$$

بالتوقع . وهذا المقدار سيكون ثابتاً (لانه تكامل محدود) ولكنه يعتمد على الثوابت التي تتضمنها دالة كثافة الاحتمال f(x)

وإذا كان المتغير x متقطعا يصبح التوقع في هذه الحالة يساوي

$$\sum_{\text{all } x} x\, P(x)$$

حيث P(x) الدالة الاحتمالية للمتغير x

سنرمز لتوقع x بالرمز E(X) حيث الحرف E يشير إلى عملية رياضية معينة تجرى على المتغير x ومضمونها :

$$E(x) = \int_{-\infty}^{\infty} x f(x)\,dx = \mu$$

أو

$$= \sum_{\text{all } x} xP(x) \quad = \mu$$

حسب كون المتغير x متصلا أو متقطعا.

كما يدعى التوقع E(x) بالوسط الحسابي للتوزيع الاحتمالي ويستخدم الحرف الإغريقي μ (ميو) في الغالب للتعبير عن القيمة المتوقعة للمتغير x .

ويلاحظ من صيغة التوقع (في حالة المتغير المتقطع مثلا) أنها تشبه أوزانا P(x) معلقة عند النقط x_i (i = ٠ ، ١ ، ٢ ، ...) والتوقع يقابل حساب مركز ثقل هذه الأوزان وهي نقطه الاتزان .

إذا كان x متغيرا عشوائيا فان أي دالة في x (لنقل مثلاً $\varphi(x)$) هي الأخرى متغير عشوائي وتوقعه يساوي:

$$E[\varphi(x)] = \sum_{\text{all } x} \varphi(x) P(x)$$

$$= \int_{\substack{\text{Range} \\ x}} \varphi(x)\, f(x)\, dx$$

حسب كون x متقطعا أو متصلا .

لنأخـ ـذ $\varphi(x) = x^r$

يكـ ـ ون

$$E(x^r) = \sum_{\text{all } x} x^r\, P(x)$$

$$= \int_{Rx} x^r f(x)\, dx$$

ولنجعل $r = 1$

يكـ ـون

$$E(X) = \sum_{\text{all } x} x \; P(x) = \mu$$

$$= \int_{Rx} x\, f(x)\, dx = \mu$$

مثـ ـال ١ :

ألقيت قطعة نقد معدنية ثلاث مرات متتالية فما هو التوقع الرياضي لعدد مرات ظهور "صورة"؟

الحـ ـل

يمكننا أن نمثل الفراغ العيني والاحتمالات المناظرة في الجدول التالي:

الحالات الممكنة	عدد ظهور الصورة	الاحتـ ـمال
ص ص ص	٣	١/٨
ص ص ك	٢	١/٨
ص ك ص	٢	١/٨
ص ك ك	١	١/٨
ك ك ك	٠	١/٨
ك ك ص	١	١/٨
ك ص ك	١	١/٨
ك ص ص	٢	١/٨

▪ **خواص التوقـ ـع Properties of Expected Value**

عبارة عن رمز يشير إلى عملية رياضية معينه Eرأينا أن دليل التوقع تجري على المتغير الإحصائي التالي لهذا الدليل ويمتاز هذا الدليل بالخواص التالية:

إذا كان a ثابت و x متغير عشوائي فإن

$$E(ax) = aE(X)$$

البرهـ ـ ـان:

$$E(ax) = \sum_{\text{all } x} ax\,P(x)$$

$$= a\sum_{\text{all } x} X\,P(x)$$

$$= a\,E(X)$$

أو

$$E(ax) = \int_{-\infty}^{\infty} a\,xf(x)\,dx$$

$$= a\int_{-\infty}^{\infty} f(x)\,dx$$

$$= a\,E(x)$$

توقع الثابت يساوي الثابت

وإذا كانa ثابت X فإن متغير عشوائي

$$E(a) = a$$

البرهـ ـ ـان :

$$E(a) = \sum_{\text{all } x} aP(x)$$
$$= a \sum P(x)$$
$$= a(1)$$
$$= a$$

نتيجـ ـ ـة :

$$E(ax+b) = a\,E(x) + b$$

البرهـ ـ ـان :

$$E(ax+b) = E(ax) + E(b)$$
$$= a\,E(x) + b$$

مثـ ـال ٢ :

$$E(2x + 3) = 2\,E(x) + 3$$

توقع التوقع يساوي التوقع

$$E[E(x)] = E(x)$$

البرهـ ـ ـان:

بما أن التوقع قيمة ثابتة كما ذكرنا وبما أن توقع الثابت يساوي الثابت

أذن توقع التوقع يساوي التوقع.

فلو رمزنا إلى توقع $E(x)$ بالقيمة الثابتة z

فان $E(z) = z$

ومنها $E[E(x)] = E(x)$

- ٤ $E[(X - E(x)] = 0$

البرهـ ـ ـان :

$$E[(x - E(x) = E(x) - E(E(x)]$$
$$= E(x) - E(x)$$
$$= 0$$

نظريـ ـة ١ : إذا كان X و Y متغيرين عشوائيين فان :

$$E(X \pm Y) = E(x) \pm E(Y)$$

أي أن توقع مجموع متغيرين يساوي مجموع توقعات المتغيرين

البرهـ ـان:

$$E(X \pm Y) = \int_{-\infty}^{\infty}\int (x \pm y) f(x.y) d\,xdy$$
$$= \iint x f(x, y)\, dxdy \pm \iint y f(x, y) dxdy$$
$$= \int x [\int f(x, y)dy]dx \pm \int y [\int f(x, y) dx]dy$$
$$= \int x f(x) dx \pm \int y f(y)\, dy$$
$$= E(X) \pm E(y)$$

نظريـ ـة ٢ : إذا كان X , Y متغيرين عشوائيين مستقلين فان :

$$E(XY) = E(X) E(Y)$$

البرهـ ـ ـان :

$$E(X Y) = \iint X Y f(x, y)\, d x d y$$

$$E(x^2) = \int_0^1 x^2 f(x)\, dx$$

$$= \int_0^1 x^3 (2x)\, dx$$

$$= \int_0^1 2x^3\, dx$$

$$= 2\left|\frac{x^4}{4}\right|_0^1$$

$$= \frac{1}{2}$$

$$E(x+1)^2 = E[x^2 + 2x + 1]$$

$$= E(x^2) + 2E(x) + 1$$

$$= ½ + 2(2/3) + 1$$

$$= ½ + 4/3 + 1$$

$$= \frac{17}{6}$$

- **التباين والانحراف المعياري Variance & Standard Deviation**

نعرف ان التبـاين والانحـراف المعيـاري هـي مقـاييس للتشـتت (Dispersion) أو لتوزيع تكراري ولكننا سنناقشهما الآن كمقاييس للتشـتت للتوزيع الاحتمالي لمتغير عشوائي .

أن القيمة المتوقعه (أو الوسط الحسـابي) لتوزيـع احتمـالي لمتغيـر عشوائي ما تدلنا على مركز التوزيع الاحتمالي وتعطينا معلومات سريعة عـن

الوسط في المدى البعيد إذا ما استمر في إجراء تجربة ما أكثر ، ولكنها لا تعطينا أي معلومات عن مدى لنتشار أو تشتت قيم المتغير العشوائي من تجربه إلى أخرى. وسنتطرق هنا إلى أكثر المقاييس شيوعاً واستعمالا مـن مقـاييس تشتت التوزيع الاحتمالي للمتغير العشوائي وهي التباين والانحراف المعياري.

ان تباين التوزيع الاحتمالي لمتغير عشوائي يقاس غالبا بنفس طريقـة قياسه في حالة التوزيع التكراري والفرق الوحيد هو أننا الآن نجد متوسط مربع انحرافات القيم عن القيمة المتوقعة بدلا من إيجاد متوسط مربع الانحرافات عن الوسط الحسابي للعينة أو التوزيع التكراري.

- التبايـ ـن

هو متوسط مربع Xتباين التوزيع الاحتمالي لمتغير عشوائي

X عن القيمة المتوقعة للتوزيع الاحتمالي للمتغير Xانحرافات قيم

$E(x) = \mu$ فإذا كان

فان $Var(x)$ أو σ_x^2 بـ Xوإذا رمزنا الى تباين

$$\sigma_x^2 = Var(x) = \sum_{all\ x} (x-\mu)^2\, p(x)$$

$$= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\, dx$$

حسب كون المتغير متقطعا أو متصلا

$$= E[(x-u)^2]$$

$$= E[x - E(x)]^2$$

أي أن التباين هو عبارة عن توقع مربع انحرافات القيم عن توقعها.

وهكذا X الداخلة في حساب التباين هي القيم المربعة لـ Xنلاحظ أن قيم وإنما بقيمها المربعة وكي نحصل على Xفان التباين لا يقيس التشتت بنفس قيم فأننا نجد الجذر التربيعي الموجب للتباين وهو ما Xمقياس للتشتت بنفس قيم يدعى بالانحراف المعياري.

- <u>الانحراف المعياري</u>

هو الجذر التربيعي X الانحراف المعياري لمتغير عشوائي
ويرمز له بالرمز σ_x ويساوي: Xالموجب لتباين

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{E(x-\mu)^2}$$

وهكذا فأن $\sigma_x^2 = \frac{3}{4}$ التباين

الانحراف المعياري $\sigma_x = \sqrt{\frac{3}{4}}$

$$= \frac{\sqrt{3}}{2}$$

نظريـ ـة: إذا كان X متغيرـا عشـوائيا توقعـه $E(x) = \mu$ وتباينـه

فإن $Var(x) = \sigma_X^2$

$$\sigma^2 = E(x^2) - [E(x)]^2$$

$$= E(x^2) - \mu^2$$

البرهـ ـ ـان :

$$\sigma^2 = E[(x - E(x)]^2$$

$$= E[(x^2 - 2xE(x) + \{E(x)\}^2]$$

$$= E(x^2) - 2\{E(x)\}^2 + \{E(x)\}^2$$

$$= E(x2) - [E(x)]^2$$

▪ خـ ـواص التبايـ ـن Properties of Variance

هناك عده خواص للتباين وهذه الخواص تسهل علينا حساب التباين لأي داله في المتغير العشوائي X وأهم الخواص هي:

١) إذا كان **a** ثابت و **X** متغير عشوائي فإن

$$Var\,(ax) = a^2\,Var\,(x)$$

البرهـ ـ ـان :

$$Var\,(ax) = E\,[ax - E\,(ax)]^2$$

$$= E\,[ax - a\,E\,(x)]^2$$

$$= a^2E\,[x - E\,(x)]^2$$

$$= a^2 var\,(x)$$

<u>مثـ ـال ٦</u> :

إذا كان تباين x يساوي ٠.٥ فما هو تباين المتغيرات التالية:

i) $2\,X$

ii) $\dfrac{X}{2}$

<u>الحـ ـل</u>:

i) $$Var\,(2x) = 4\,Var\,(x)$$

$$= 4\,(0.5)$$

$$= 2$$

ii) $$\text{Var}\left(\frac{X}{2}\right) = \frac{1}{4}\ \text{Var}(x)$$

$$= \frac{1}{4}\ (0.5)$$

$$= 0.125$$

٢) تباين الثابت يساوي صفر

إذا كان a ثابت فان

$$\text{Var}\ (a) = 0$$

البرهـ ـ ـان :

$$\text{Var}(a) = E\,[a - E(a)]^2$$

$$= E\,(a - a)^2$$

$$= E\,(0^2)$$

$$= 0$$

نتيجـ ـ ـة :

$$\text{Var}(x \pm a) = \text{Var}(x)$$

البرهـ ـ ـان :

$$\text{Var}(x \pm a) = \text{Var}(x) \pm \text{Var}(a)$$

$$= \text{Var}(x) \pm 0$$

$$= \text{Var}(x)$$

وهكذا فان إضافة أي قيمة أو طرح أي قيمة ثابته إلى أو من المتغيـ ـر العشوائي لا تؤثر على التباين لهذا المتغير.