الفصل الأول

مفاهيم أساسية في علم الإحصاء والاحتمال

علم الإحصاء:

مجموعة من الطرق الرياضية التي تستخدم لجمع وتبويب وتنظيم وتلخيص وعرض المعلومات والبيانات ومن ثم تحليلها وفق طرق وأساليب علمية للحصول على استدلالات وقرارات سليمة.

فروع علم الإحصاء:

أولاً: الإحصاء الوصفي:

هو الفرع المتعلق بطرق جمع وتبويب وتنظيم وتلخيص وعرض المعلومات والبيانات ووصف توزيع البيانات

ثانياً: الإحصاء الاستدلالي:

هو الفرع المتعلق بالطرق المساعدة في إتخاذ قرارات حول مجتمع وذلك عن طريق دراسة عينة إحصائية من ذلك المجتمع.

المجتمع والعينة:

أولاً: المجتمع:

هو المجموعة الكلية من الأشياء (تسمى عناصر المجتمع) ذات خصائص مشتركة ويكون الباحث مهتماً بها حيث سيتم عمل بعض الاستدلالات والنتائج حولها.

ينقسم المجتمع إلى:

1. **المجتمع المحدود:** عدد عناصر هذا المجتمع يكون محدوداً أو منتهياً مثل طلاب كلية العلوم بجامعة الملك سعود.
2. **المجتمع غير المحدود:** عدد عناصر هذا المجتمع يكون غير محدود أو غير منته مثل المجتمع المكون من النجوم في السماء.

ثانياً: العينة:

هي مجموعة مكونة من عدد من أفراد المجتمع يتم اختيارهم بطريقة مناسبة بحيث تمثل المجتمع تمثلاً جيداً وذلك لدراسة صفات المجتمع إذ يستدل بصفات العينة على صفات المجتمع.

حجم العينة: يعرف حجم العينة على أنه عدد أفراد العينة ويرمز له عادة بالرمز n.

المتغيرات:

المتغير هو مقدار كمي أو وصفي ويستخدم لقياس خاصية معينة لعناصر مجتمع أو عينة والتي تختلف من عنصر لآخر.

ومن أمثلة المتغيرات: طول شخص، فصيلة دم لشخص، والمستوى التعليمي لشخص.

أنواع المتغيرات: يمكن تصنيف المتغيرات إلى فئتين رئيسيتين على النحو التالي:

1. متغير كمي:

إن قيم المتغير الكمي عبارة عن أرقام أو قيم عددية تدل على كميات أو أعداد أشياء معينة لأفراد المجتمع.

ومن أمثلة المتغيرات الكمية ما يلي: طول شخص وعدد أولاد شخص.

2. متغير وصفي (أو نوعي):

إن قيم المتغير الوصفي (أو النوعي) عبارة عن صفات أو كلمات تدل فقط على انتماء أفراد المجتمع إلى فئات أو أصناف معينة.

ومن أمثلة المتغيرات الوصفية ما يلي: فصيلة دم لشخص والمستوى التعليمي لشخص.

البيانات:

البيانات هي مجموعة المشاهدات أو الملاحظات المأخوذة من الدراسة الإحصائية.

وتنقسم البيانات إلى نوعين تبعاً لنوع المتغير:

1. **بيانات كمية** (مثل: بيانات أطوال مجموعة من الأشخاص)
2. **بيانات وصفية** (مثل: بيانات المستوى التعليمي لمجموعة من الأشخاص)

ومن أمثلة المتغيرات الوصفية ما يلي: فصيلة دم لشخص والمستوى التعليمي لشخص.

المعلمة والإحصاءة:

المعلمة:

هي مقدار (أو خاصية) يتميز بها المجتمع (مثل متوسط أعمار طلاب كلية العلوم). وتحسب العملمة من المجتمع ولذلك فهي ممقدار مجهول في الغالب ونرغب في معرفته.

الإحصاءة:

مقدار (أو خاصية) تتميز بها العينة (مثل متوسط أعمار عشرة طلاب من طلاب كلية العلوم). وتحسب الإحصاءة من العينة ولذلك فهي مقدار معلوم.

الفصل الثاني

تنظيم البيانات وعرضها

تنظيم البيانات وتلخيصها وعرضها جدولياً:

بعد أن يتم تحديد مجتمع الدراسة والعينة وقياس قيم المتغير على أفراد العينة نحصل على البيانات في صورتها الأولية والتي تسمى **البيانات الخام**.

إن البيانات في صورتها الخام تحتاج إلى تنظيم وتلخيص لكي نتمكن من عرضها بطريقة مناسبة يمكن الاستفادة منها.

ولتصميم جدول التوزيع التكراري للبيانات يلزمنا التأكيد على أن البيانات تنقسم إلى نوعين هما:

1. **بيانات وصفية:** مثل لون الشعر، فصيلة الدم، الجنس، المستوى التعليمي، وغيرها.
2. **بيانات كمية:** مثل الطول، الوزن، العمر، عدد الأولاد، وغيرها.

والبيانات الإحصائية سواءً كانت وصفية أو كمية فهي تنظم وتلخص في جداول تسمى الجداول التكرارية.
والجدول التكراري عبارة عن جدول يلخص البيانات فيوزعها على فئات (أو فترات) ويحدد عدد البيانات التي تنتمي لكل فئة (تكرار الفئة).

الجداول التكرارية للبيانات الوصفية:

أولاً: الجدول التكراري للبيانات الوصفية:

1. نحصر جميع الفئات وهي الصفات المختلفة في البيانات.
2. ننشئ جدول تفريغ البيانات وهو عبارة عن جدول مكون من ثلاثة أعمدة هي: عمود الفئة (أو الصفة) وعمود العلامات و عمود التكرار.
3. نوجد الجدول التكراري من جدول تفريغ البيانات وذلك بحذف عمود العلامات.

مثال 1: أوجد الجدول التكراري لتقديرات عينة من ستين طالباً معطاة فيما يلي:

D	B	E	C	D	B	D	C	E	A	D	B	C	C	C
B	E	C	D	B	D	D	A	E	C	A	D	E	C	C
C	D	A	C	E	D	C	C	D	B	B	C	D	E	D
D	E	D	D	A	D	D	C	D	C	D	D	B	D	A

الحل:

الفئة أو الصفة (التقدير)	التكرار

ثانياً: الجدول التكراري النسبي والجدول التكراري المئوي للبيانات الوصفية:

نهتم كثيراً بمعرفة التكرارات النسبية والتكرارات المئوية لكل فئة من فئات الجدول التكراراي. فمعرفتنا بالتكرار النسبي للفئة أكثر فائدة من مجرد معرفة التكرار.

نعرف التكرار النسبي والتكرار المئوي للفئة كما يلي:

1. **التكرار النسبي للفئة = تكرار الفئة \ عدد البيانات**

= تكرار الفئة \ مجموع التكرارات

2. **التكرار المئوي للفترة = التكرار النسبي x 100%**

مثال 2: أوجد الجدول التكراري النسبي والجدول التكراري المئوي في مثال (1).

الفئة أو الصفة (التقدير)	التكرار (f)	التكرار النسبي= f/n	التكرار المئوي= 100%*(f/n)
A	6		
B	8		
C	16		
D	22		
E	8		
المجموع	n=60		

الفصل الثالث
مقاييس النزعة المركزية

تنقسم المقاييس الإحصائية إلى قسمين:

- مقاييس النزعة المركزية.
- مقاييس التشتت.

أولاً: مقايسس النزعة المركزية

الوسط الحسابي | الوسيط | المنوال

1) الوسط الحسابي

- وهو من أهم مقاييس النزعة المركزية والأكثر استخداماً في الإحصاء و في الحياة العملية.

- **تعريف:**

- إذا كان لدينا مجموعة من المشاهدات للمتغير X وهي $x_1, x_2, \dots, x_n$ فإن الوسط الحسابي (ويرمز له $\bar{X}$) يساوي مجموع المشاهدات مقسوماً على عددها.

- $$\bar{X} = \frac{\sum_{i=1}^{n} x_i}{n}$$

مثال

أوجدي المتوسط الحسابي للمشاهدات التالية والتي هي عبارة عن أوزان (بالكيلو جرام)

25 30 40 45 35 55 50

الحل:

$$\bar{x} = \frac{\sum x_i}{n} = \frac{25 + 30 + \cdots + 50}{7} = 40 \text{ كيلوجرام}$$

- مثال
- إذا كان المتوسط الحسابي للمشاهدات $x_1, x_2, \ldots\ldots x_n$ هو 15 فإن المتوسط الحسابي للمشاهدات $\frac{x_1 - 10}{2}, \frac{x_2 - 10}{2}, \ldots\ldots \frac{x_n - 10}{2}$

هو ....

- **خصائص الوسط الحسابي**
- المجموع الجبري لإنحرافات القيم عن وسطها الحسابي يساوي صفر.

$$\sum_{i=1}^{n}(\mathrm{x_i} - \bar{\mathrm{x}}) = 0$$

- إذا جمعنا أو طرحنا من المشاهدات الأصلية جميعها قيمة ثابتة ولتكن a فإن الوسط الحسابي للمشاهدات بعد الجمع يساوي الوسط الحسابي قبل الجمع مضافا إليه الثابت a .
- أي أن
- الوسط الحسابي (بعد الجمع) = الوسط الحسابي (قبل الجمع) +a
- الوسط الحسابي (بعد الطرح) = الوسط الحسابي (قبل الطرح) -a
- إذا ضربنا أو قسمنا من المشاهدات الأصلية جميعها قيمة ثابتة ولتكن b فإن الوسط الحسابي للمشاهدات بعد الضرب يساوي الوسط الحسابي قبل الضرب مضروباً في الثابت b.
- أي أن:
- الوسط الحسابي (بعد الضرب) = الوسط الحسابي (قبل الضرب) *b
- الوسط الحسابي (بعد القسمة) = الوسط الحسابي (قبل القسمة) \ b

مميزات الوسط الحسابي

- مقايس سهل ويخضع للعمليات الجبريه بسهوله.
- يأخذ في الإعتبار جميع القيم محل الدراسة.
- المتوسط وحيد لمجموعة البيانات الواحدة

عيوب الوسط الحسابي

- يتأثر بالقيم المتطرفة (وهي القيم الكبيرة جداً أو الصغيرة جداً)
- يصعب حسابه في حالة البيانات الوصفية.

المتوسط المرجح:

في بعض الأحيان تكون المشاهدات $x_1, x_2, \ldots\ldots x_n$ مقرونة بالأوزان $w_1, w_2, \ldots\ldots w_n$ على التوالي وفي هذه الحالة نعرف الوسط المرجح

$$\overline{x_w} = \frac{\sum x\,w}{\sum w} = \frac{x_1 w_1 + x_2 w_2 + \cdots \ldots + x_n w_n}{w_1 + w_2 + \cdots + w_n}$$

- مثال

أوجد المتوسط المرجح لدرجات أحد الطلاب في مقررات الإحصاء والفيزياء والرياضيات باعتبار أن الوزن هو عدد الساعات للمقرر كما يلي

المقرر	عدد الساعات w	الدرجة x
احصاء	2	40
فيزياء	4	65
رياضيات	3	70

الحل :

2)الوسيط:

عند ترتيب البيانات تصاعدياً أو تنازلياً يكون الوسيط هي القيمة التي تقع قبلها 50% من المشاهدات في الترتيب و 50% من المشاهدات بعدها في الترتيب أي أنها القيمة التي تقسم البيانات إلى قسمين متساويين (المشاهدة التي تقع المنتصف).

- مثال:
- أوجدي الوسيط لمجموعة الأوزان (بالكيلو جرام) التالية

7.1 2.5 2.5 5.4 8.3

مثال:

أوجدي الوسيط لمجموعة الأوزان (بالكيلو جرام) التالية

8.3 9.2 5.4 2.5 2.5 7.1

مميزات الوسيط

1. لا يتأثر بالقيم المتطرفة.
2. يمكن حسابه في حالة البيانات الوصفية التي يمكن ترتيبها.
3. الوسيط وحيد لمجموعة البيانات الواحدة.
4. الوسيط سهل التعريف والحساب

عيوب الوسيط

1. لا يأخذ جميع القيم في الإعتبار عند حسابه.
2. لا يمكن حساب الوسيط للبيانات الوصفية الغير قابلة للترتيب

3)المنوال

- وهو القيمة الأكثر تكراراً بين البيانات (المشاهدات).

ويتقسم المنوال إلى:

عديم المنوال
(أي لا يوجد منوال)

وحيد المنوال (أي يوجد منوال واحد فقط)

متعدد المنوال
(أي يوجد أكثر من منوال)

- مثال

الجدول التالي يتضمن بيانات إحدى الدراسات التي طبقت على خمسة أشخاص لقياس العمر بالسنة والوزن بالكيلوجرام والطول بالسم وفصيلة الدم .أوجد المنوال للبيانات المختلفة:

رقم الشخص	1	2	3	4	5
العمر	25	20	25	30	35
الوزن	70	55	65	70	65
الطول	164	162	155	165	158
فصيلة الدم	O	A	B	A	AB

مميزات المنوال

- هو مقياس من السهل حسابه.
- اقل تأثر من المتوسط بالقيم الشاذة.
- يمكن إيجاده للبيانات الكمية والوصفية .

عيوب المنوال

- لا يأخذ جميع القيم في الإعتبار عند حسابه.
- قد يكون لبعض البيانات أكثر من منوال وبذلك لا يمكن تحديد قيمة واحدة للمنوال.
- بعض البيانات لا يكون لها منوال.

الربيعات والعشيرات والمئينات

- **حيث هناك ثلاث ربيعات (Quartiles) وهي كالتالي**
- Q_1 : الربيع الأول : القيمة التي تكون فيها 25% من البيانات أقل منها (قبلها) و 75% من البيانات أكبر منها (بعدها).
- Q_2 :الربيع الثاني : القيمة التي تكون فيها 50% من البيانات أقل منها (قبلها) و 50% من البيانات أكبر منها (بعدها) وهو يساوي الوسيط.
- Q_3 :الربيع الثالث : القيمة التي تكون فيها 75% من البيانات أقل منها (قبلها) و 25% من البيانات أكبر منها (بعدها).

- D1= العشير الأول
- = القيمة التي يسبقها $\frac{1}{10}$ من البيانات أو 10% من البيانات
- D2= العشير الثاني
- = القيمة التي يسبقها $\frac{2}{10}$ من البيانات أو 20% من البيانات

..............وهكذا

- D9= العشير التااسع
- = القيمة التي يسبقها $\frac{9}{10}$ من البيانات أو 90% من البيانات

- P1= المئين الأول
- = القيمة التي يسبقها $\frac{1}{100}$ من البيانات أو1 % من البيانات
- P2= المئين الثاني
- = القيمة التي يسبقها $\frac{2}{100}$ من البيانات أو2 % من البيانات

................وهكذا

- P99= المئين االتاسع و التسعين
- = القيمة التي يسبقها $\frac{99}{100}$ من البيانات أو99% من البيانات

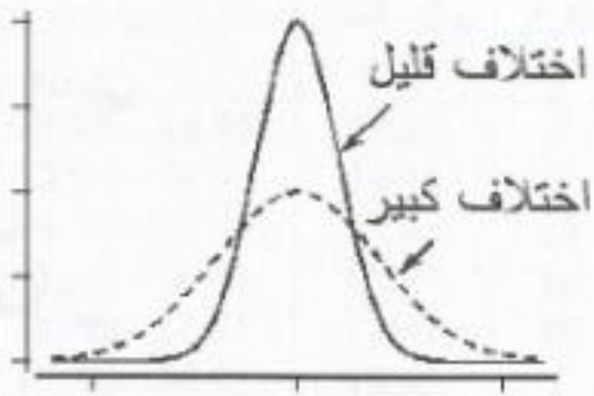
الفصل الرابع

مقاييس التشتت

مثال 1: الجدول التالي يوضح مجموعتين لهما نفس المتوسط $\bar{x} = 60$ ونفس الوسيط Med=60 ونفس المنوال Mod=60 ولكنهما مختلفان في طبيعة التشتت.

المجموعة	البيانات	شكل انتشار البيانات
الأولى	59,61,62,60,58,60	• …. • …. …. •….. …. … • …. …. …. • …. …… • …. **60**
الثانية	50,60,66,54,60,70	• ……. • ……. ……. •…… ……. …… ……. • …… …… ….. ….. • ……. ……. •……. **60**

مثال 2: الشكل التالي يوضح المضلعين التكرارين لمجموعتين من البيانات لهما نفس مقاييس النزعة المركزية ولكنهما مختلفان في طبيعة التشتت.



شكل (1) المضلعان التكراريان لتوزيعين لهما نفس مقاييس النزعة المركزية ولكنهما مختلفين في التشتت

مقاييس التشتت:

- المدى Range
- نصف المدى الربيعي Semi-Inter-quartile Range
- التباين Variance
- الانحراف المعياري Standard deviation
- معامل الاختلاف(التغير) Coefficient of Variation

المدى (Range):

يعرف المدى لمجموعة من البيانات أنه الفرق بين أكبر قيمه و أصغر قيمه أي

$$Range = X_{max} - X_{min}$$

مثال 3: أوجدي المدى للمشاهدات التالية والتي هي عبارة عن أوزان (بالكيلوجرام) مجموعه من مكونه من سبعة أشخاص 50 55 35 45 40 30 25.

مميزات وعيوب المدى:

- يتميز المدى بسهولة التعريف والحساب .
- يعيب المدى العيوب التاليه :

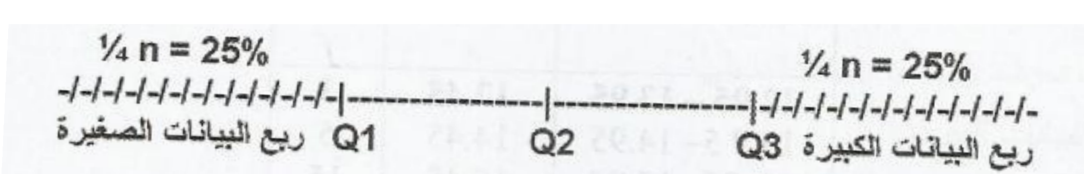
1. يتأثر المدى بالقيم الشاذة أو المتطرفه
2. لا يأخذ المدى في الاعتبار جميع البيانات

ملاحظة:

1- وحدة المدى هي نفس وحدة البيانات الأصلية.

2- نظرا لان المدى يعتمد فقط على أكبر وأصغر قيمة ولايأخذ في الاعتبار القيم الأخرى فهو مقياس غير جيد لقياس التشتت.

نصف المدى الربيعي:



ويعرف بالصيغة التالية:

$$Q = \frac{Q_3 - Q_1}{2}$$

حيث Q_1 هو الربيع الأول و Q_3 هو الربيع الثالث.

مميزات و عيوب نصف المدى الربيعي :

1- من المميزات أنه لايتأثر بالقيم الشاذة أو المتطرفة.

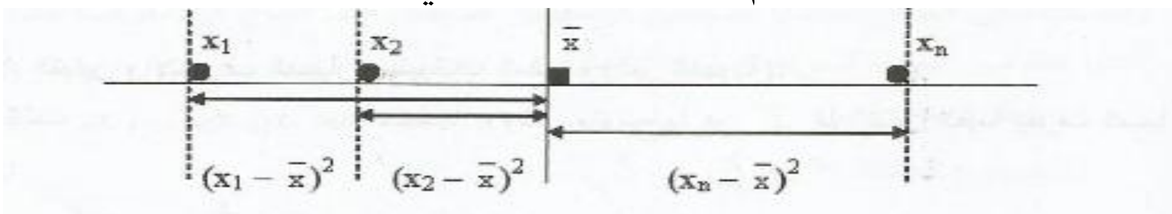
2- من العيوب أنه لاياخذ في الاعتبار جميع البيانات .

ملاحظة: وحدة نصف المدى الربيعي هي نفس وحدة البيانات الأصلية.

التباين:

يعرف التباين أنه متوسط مربع انحرافات القيم عن الوسط الحسابي ويرمز له S^2

الشكل التالي يبين مربعات انحرافات القيم عن الوسط الحسابي.



والجدول التالي يلخص طريقة حساب انحرافات ومربعات انحرافات القيم عن المتوسط.

القيم (البيانات)	x_n	$\cdots$	x_2	x_1
انحرافات القيم عن المتوسط	$x_n - \bar{x}$	$\cdots$	$x_2 - \bar{x}$	$x_1 - \bar{x}$
مربع انحرافات القيم عن المتوسط	$(x_n - \bar{x})^2$	$\cdots$	$(x_2 - \bar{x})^2$	$(x_1 - \bar{x})^2$

أي قانون التباين يصبح

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

وهناك صيغ أخرى:

$$S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2\right)$$

حيث الصيغة الثانية تستخدم أكثر لسهولتها ولأنها أكثر دقة.

الانحراف المعياري:

الجذر التربيعي الموجب للتباين ويرمز له بالرمز S .

$$S = \sqrt{S^2} = \sqrt{\frac{\sum (x_i - \bar{x})^2}{n-1}}$$

ملاحظة:

1. $S^2 \geq 0$ (دائما) وكذلك $S \geq 0$ (دائما).
2. $S^2 = 0 \Leftrightarrow S = 0 \Leftrightarrow$ جميع قيم العينه متساوية.
3. وحدة التباين هي نفس وحدة البيانات الأصلية مربعة.
4. وحدة الانحراف المعياري هي نفس وحدة البيانات الأصلية.

بعض خصائص التباين والانحراف المعياري

الانحراف	التباين	المشاهدات
S	s^2	$x_1, x_2, \dots, x_n$
S	s^2	$x_1 \pm b, x_2 \pm b, \dots, x_n \pm b$
$\lvert a \rvert s$	$a^2 s^2$	$ax_1, ax_2, \dots, ax_n$
$\lvert a \rvert s$	$a^2 s^2$	$ax_1 \pm b, ax_2 \pm b, \dots, ax_n \pm b$

مثال 6: أوجدي تباين وانحراف العينة لمجموعة الأوزان (بالكيلوجرام) التالية:

7.1, 2.5 , 2.5, 5.4, 8.3

أولا: نحسب المتوسط

$$\bar{X} = \frac{\sum_{i=1}^{5} x_i}{5} = \frac{25.8}{5} = 5.16$$

نستخدم القانون $S^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} x_i^2 - n(\bar{x})^2\right)$ لحساب التباين

$$S^2 = \frac{1}{5-1}\left(160.96 - 5(5.16)^2\right)$$

$$S^2 = 6.958$$

حساب الانحراف $S = \sqrt{6.958} = 2.6378$

كما يمكن استخدام الاله الحاسبه..

x_i	x_i^2
7.1	50.41
2.5	6.25
2.5	6.25
5.4	29.16
8.3	68.89
Total= 25.8	160.96

مقاييس التشتت التي ذكرت سابقا تكون غير مناسبة لمقارنة تجانس مجموعات البيانات المختلفة في الحالتين التاليتين:

1- إذا كانت وحدتا المتغيرين مختلفين حيث لايستطيع مقارنة الوحدات المختلفة ببعض.

2- إذا كان متوسطا المتغيرين مختلفين وذلك لأن تباين توزيع المتغير ذي المتوسط الصغير ينزع لأن يكون صغيرا والعكس بالعكس.

معامل الاختلاف:

أحد مقاييس التشتت النسبي ويستخدم لمقارنة التشتت النسبي أو التجانس لمجموعة البيانات المختلفة

$$C.V = \frac{S}{\bar{x}}$$

مثال 11: الجدول أدناه يتضمن بيانات إحدى الدراسات التي طبقت على خمسة أشخاص لقياس الوزن (بالكيلوجرام) والطول(السنتيمتر) .أي البيانات أكثر تشتتا نسبا (أقل تجانسا) ؟ بيانات الأوزان أم بيانات الأطوال؟

رقم الشخص	1	2	3	4	5
الأوزان	69	59	65	67	65
الأطوال	164	162	155	165	158

نوجد المتوسط والانحراف من الاله الحاسبة.

البيانات	المتوسط $\bar{x}$	الانحراف المعياري S	معامل الاختلاف C.V
الأوزان	65.0 kg	3.7417 kg	0.0576
الأطوال	160.8 cm	4.2071 cm	0.026

بما أن معامل الاختلاف لبيانات الأوزان أكبر من معامل الاختلاف لبيانات الأطوال فإن التشتت النسبي لبيانات الأوزان أكبر من التشتت النسبي لبيانات الأطوال . أي أن بيانات أقل تجانسا من بيانات الأطوال.

نظرية تشيبيشيف:

تعطي **حد أدنى** لنسبة البيانات الواقعة في فترة معينة عند معرفة متوسط البيانات وانحرافها المعياري دون الحاجة لمعرفة البيانات الأصلية.

- **نظرية:**

إذا كان لدينا عينة من البيانات متوسطها $\bar{x}$ و انحرافها المعياري S فإن نسبة البيانات الواقعة في الفترة ($\bar{x}$-ks, $\bar{x}$+ks) **لا يقل عن**

$$\left(1-\frac{1}{k^2}\right)$$

حيث k>1.

ملاحظات:

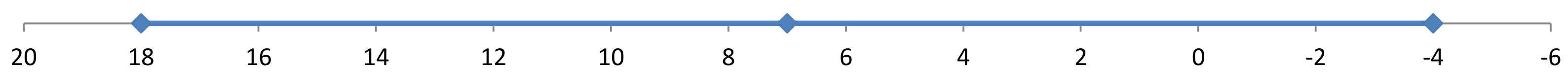
1. تطبق نظرية تشيبيشف للفترات التي منتصفها هو المتوسط.

2. تستخدم نظرية تشيبيشف بطريقتين

أ. تحديد النسبة التقريبية لعدد البيانات الواقعة في فترة معينة

ب. تحديد فترة قع فيها ما لا يقل عن نسبة معينة

مثال 12: إذا كان لدينا مجموعة من البيانات متوسطها 7 وانحرافها المعياري 5 فما هي نسبة البيانات الواقعة في الفترة (-4,18)



مثال 13: إذا كان لدينا مجموعة من البيانات متوسطها 7 وانحرافها المعياري 5 فأوجد فترة يقع فيها ما لا يقل عن 75% من البيانات.

الدرجات (القيم) المعيارية:

تستخدم لمقارنة البيانات في المجموعات المختلفة

لتكن $x_1, x_2, \dots, x_n$ عينة من البيانات حجمها n ومتوسطها $\bar{x}$ وانحرافها المعياري S. نعرف الدرجة (القيمة) المعيارية للمشاهدة x_i بالصيغة التالية:

$$z_i = \frac{x_i - \bar{x}}{s}$$

ملاحظة: القيمة $z_i = \frac{x_i - \bar{x}}{s}$ هي الدرجة المعيارية للمشاهدة الأصلية x_i

مثال 14: إذا كان لدينا مجموعة من البيانات متوسطها 7 وانحرافها المعياري 5 فأوجد:

1. **الدرجة المعيارية للقيمة x=9.**

2. **القيمة الأصلية للدرجة المعيارية z=0.1.**

مثال (15)

إذا كانت درجة أحد الطلاب في مقرر الاحصاء تساوي 82 ودرجته في مقرر الرياضيات 89 واذا كان متوسط درجات الطلاب في مقرر الرياضيات يساوي 81 بانحراف معياري يساوي 16

ففي اي المقررين كان أداء الطالب أفضل

الفصل الخامس

طرق العد

القواعد الأساسية لطرق العد:

أولاً: قاعدة الضرب:

$$n=n_1 \times n_2 \times \dots \times n_r$$

مثال 1: بكم طريقة مختلفة يمكن أن يختار أحد الطلاب ثلاثة مقررات: الأول في الإحصاء والثاني في الرياضيات والثالث في الفيزياء إذا علم أن هناك 3 مقررات مختلفة للإحصاء و مقررين مختلفين للرياضيات و مقررين مختلفين للفيزياء.

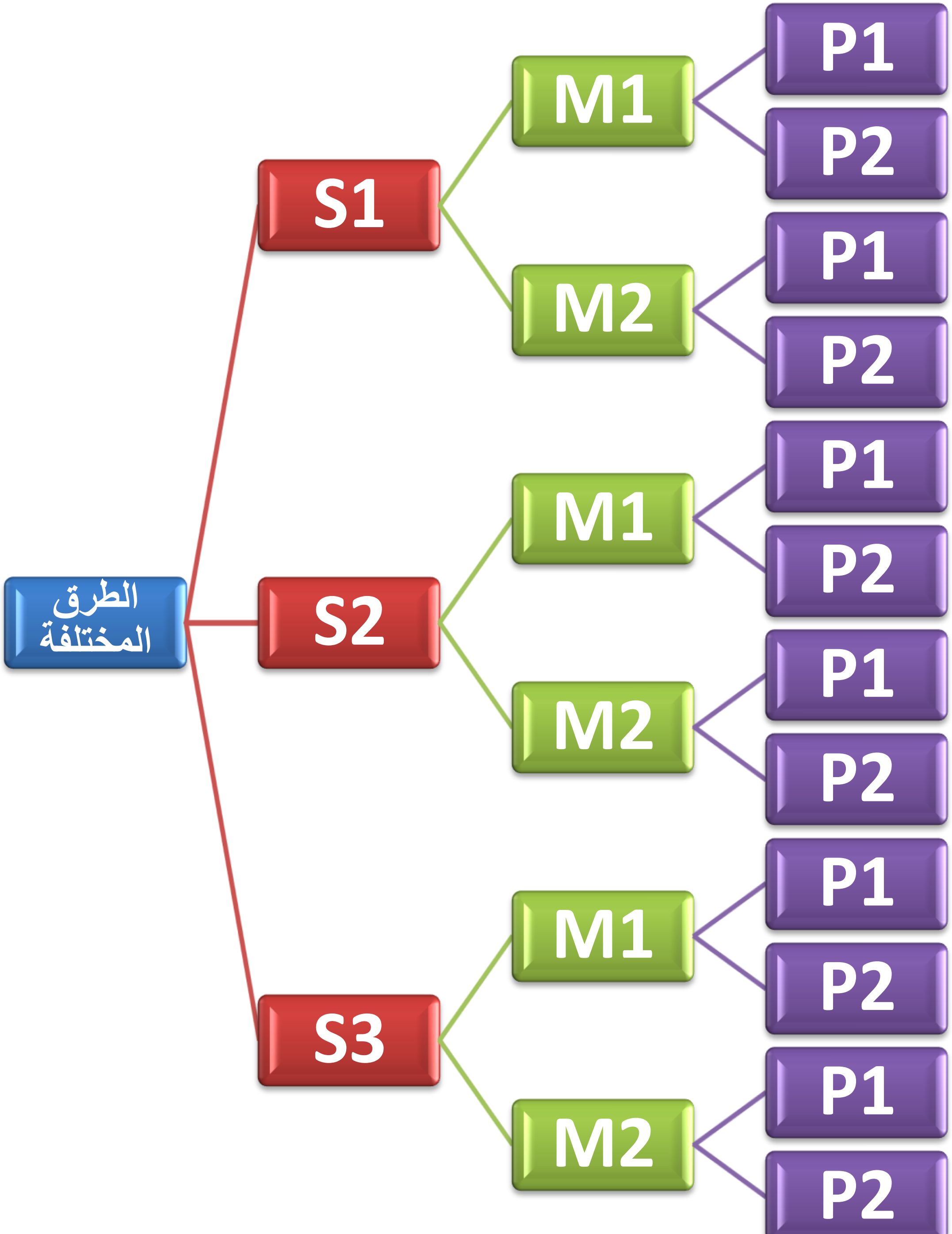
عدد طرق اختيار مقرر الأحصاء يساوي $n_1= 3$

عدد طرق اختيار مقرر الرياضيات يساوي $n_2= 2$

عدد طرق اختيار مقرر الفيزياء يساوي $n_3= 2$

باستخدام قاعدة الضرب يكون عدد الطرق المختلفة لاختيار المقررات الثلاثة يساوي

$$n= 3 \times 2 \times 2 = 12$$



الطرق المختلفة
S1
S2
S3
M1
M2
M1
M2
M1
M2
P1
P2
P1
P2
P1
P2
P1
P2
P1
P2
P1
P2

ثانياً: قاعدة الجمع:

$$n=n_1 + n_2 + \ldots + n_r$$

مثال 2: بكم طريقة مختلفة يمكن أن يختار أحد الطلاب مقرراً واحداً فقط من الإحصاء أو الرياضيات أو الفيزياء إذا علم أن هناك 3 مقررات مختلفة للإحصاء و مقررين مختلفين للرياضيات و مقررين مختلفين للفيزياء.

عدد طرق اختيار مقرر الأحصاء يساوي $n_1= 3$

عدد طرق اختيار مقرر الرياضيات يساوي $n_2= 2$

عدد طرق اختيار مقرر الفيزياء يساوي $n_3= 2$

وحيث أن العمليات متنافية وباستخدام قاعدة الجمع يكون عدد الطرق المختلفة لاختيار المقرر الثلاثة يساوي

$$n= 3 + 2 + 2 = 7$$

التباديل:

التبديلة هي ترتيبة لعدة أشياء بأخذها كلها أو بعضها في كل مرة مع مراعاة الترتيب.

مثال 3:

1. **كم عدد تباديل الحروف A, B, C بحيث تحتوي كل ترتيبة على حرفين؟ أو بعبارة أخرى بكم طريقة يمكن ترتيب حرفين من الحروف A, B, C.**

عدد طرق اختيار الحرف الأول $n1 = 3$

عدد طرق اختيار الحرف الثاني $n2 = 2$

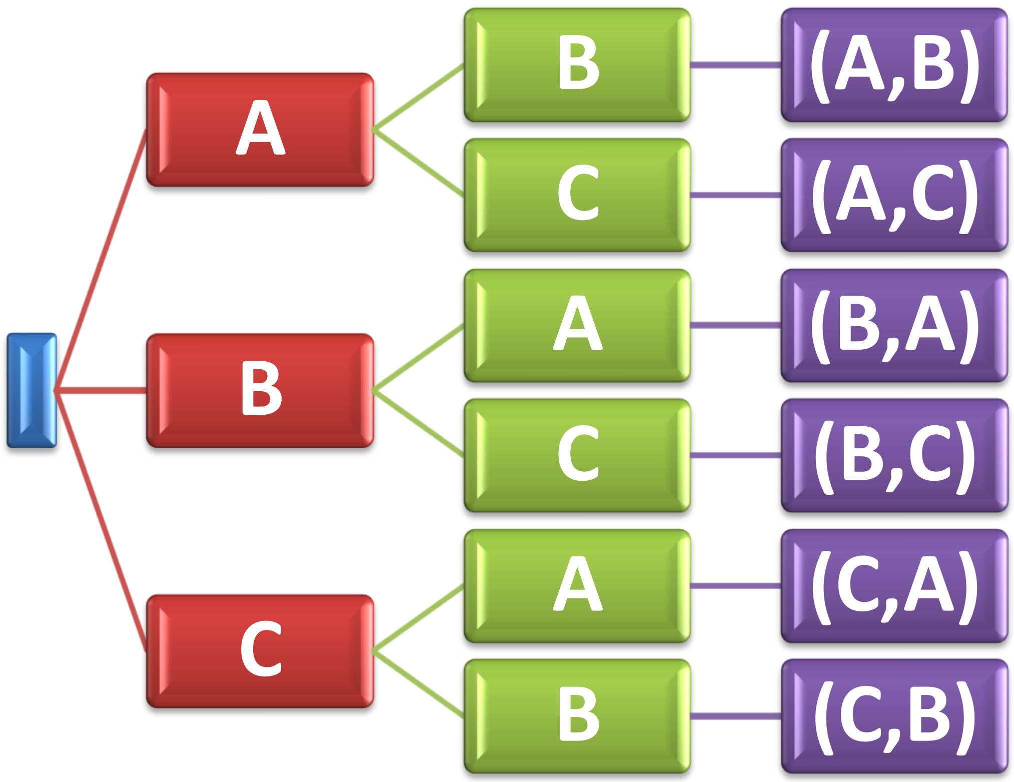
باستخدام قاعدة الضرب يكون عدد طرق ترتيب حرفين من الحروف يساوي:

$$n = n1 \times n2 = 3 \times 2 = 6$$

مثال 3:

2. أوجدي التباديل المختلفة لحرفين من الحروف A, B, C.

(A,B), (A,C), (B,A), (B,C), (C,A), (C,B)



نتيجة:

عدد تباديل n من الأشياء المختلفة مأخوذة r في كل مرة يرمز له بالرمز **nPr** ويعطى بالصيغة التالية:

$$\mathbf{nPr} = \frac{n!}{(n-r)!}$$

ملاحظات:

1. يرمز لمضروب العدد n بالرمز n! ويعرف

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 2 \times 1$$

2. $0! = 1$

4. $nPn = n!$

5. $nP1 = n$

مثال 4: أوجدي

7P3 =

5P5 =

5! =

مثال 6:

1. بكم طريقة يمكن أن نجلس 5 طلاب على 5 مقاعد في صف واحد ؟

5P5 =

❖ **عدد ترتيب n من الاشياء المختلفة في صف هو: n!**

2. بكم طريقة يمكن أن نجلس 5 طلاب على 3 مقاعد في صف واحد ؟

5P3 =

تطبيقات على التباديل:

1. السحب بإرجاع:

عدد الطرق المختلفة هو $n \times n \times \ldots \times n = n^r$

2. السحب بدون إرجاع:

عدد الطرق المختلفة هو $nPr = \dfrac{n!}{(n-r)!}$

مثال 7: بكم طريقة يمكن سحب كرتين بإرجاع من صندوق يحتوي على 15 كرة مختلفة؟

مثال 8: بكم طريقة يمكن سحب كرتين بدون إرجاع من صندوق يحتوي على 15 كرة مختلفة؟

التوافيق:

التوفيقة هي كل مجموعة يمكن اختيارها من مجموعة من عدة اشياء بأخذها كلها أو بعضها دون مراعاة الترتيب.

نتيجة:

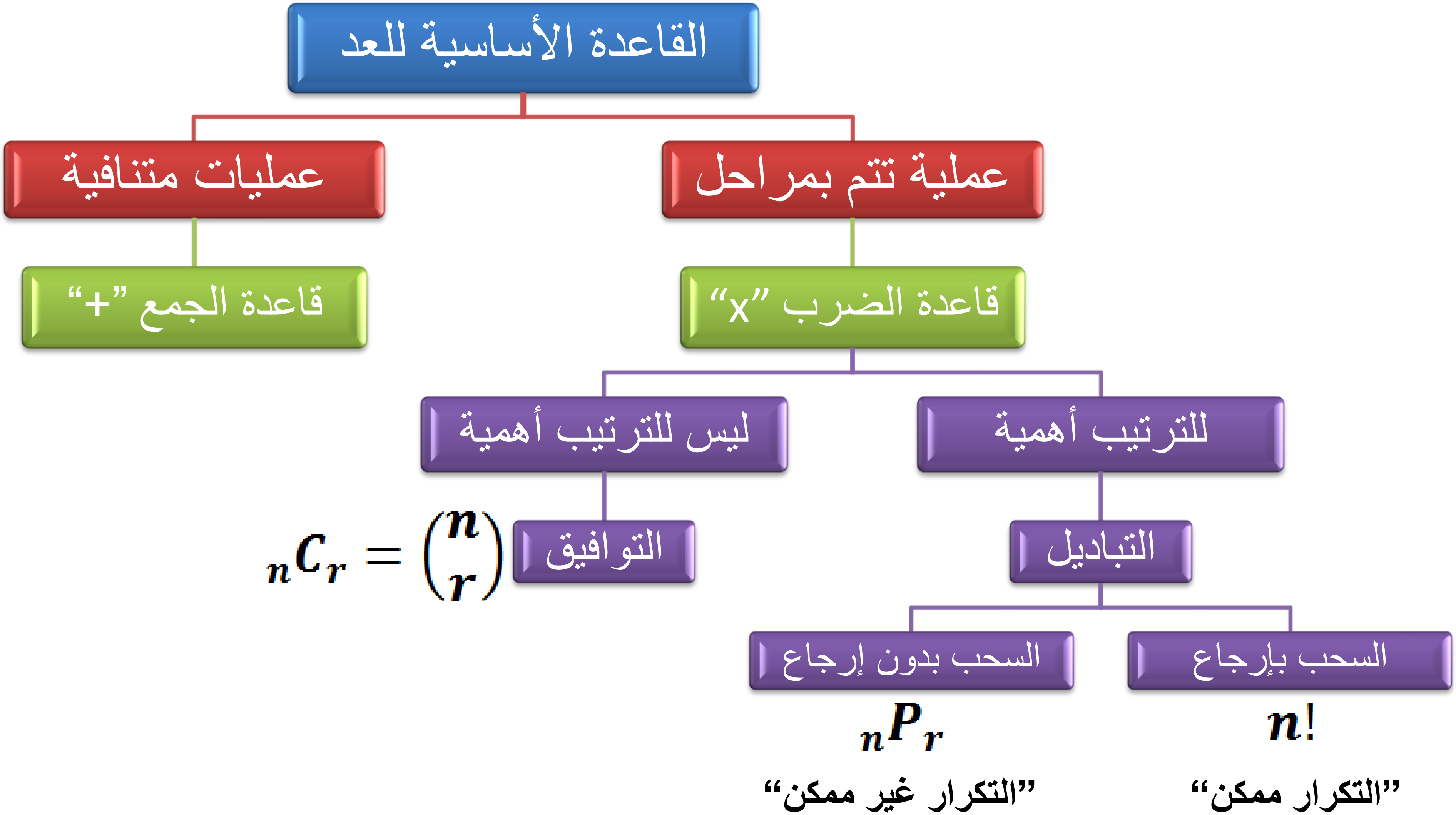
إذا كان لدينا مجموعة مكونة من n من العناصر المختلفة فإن عدد التوافيق التي يمكن تكوينها بحيث تحتوي كل توفيقة على r عنصر يرمز له بالرمز **nCr** أو $\binom{n}{r}$ ويعطى بالصيغة التالية:

$$\binom{n}{r} = \mathbf{nCr} = \frac{n!}{r!\,(n-r)!}$$

- **ملاحظة:**

1. $\binom{n}{1} = n, \binom{n}{n} = 1, \binom{n}{0} = 1$

2. $\binom{n}{r} = \binom{n}{n-r}$


القاعدة الأساسية للعد
عملية تتم بمراحل
عمليات متنافية
قاعدة الضرب "x"
قاعدة الجمع "+"
للترتيب أهمية
ليس للترتيب أهمية
التباديل
التوافيق
$_nC_r = \binom{n}{r}$
السحب بإرجاع
السحب بدون إرجاع
$n!$
$_nP_r$
"التكرار ممكن"
"التكرار غير ممكن"

مثال 9:

1. 5C3=
2. 4C0=
3. 7C7=
4. 7C1=

مثال 10: بكم طريقة يمكن اختيار حرفين (بدون مراعاة الترتيب) من مجموعة الحروف A, B, C ؟

التباديل داخل أشياء متشابهة:

إذا كان هناك n من الأشياء مكونة من r مجموعة بحيث أن المجموعة رقم 1 مكونة من n1 من العناصر المتشابهة والمجموعة رقم 2 مكونة من n2 من العناصر المتشابهة وهكذا....

$$\binom{n}{n_1, n_2, \dots, n_r} = \frac{n!}{n_1!\, n_2! \dots n_r!}$$

مثال 11: بكم طريقة يمكن ترتيب أحرف كلمة PROBABILITY.

مثال 12: بكم طريقة يمكن ترتيب أحرف كلمة STATISTICS.

مثال 13: بكم طريقة يمكن توزيع 8 طلاب على النحو التالي: 4 طلاب لتخصص الإحصاء و 3 طلاب لتخصص الرياضيات و طالب واحد لتخصص الفيزياء.

الفصل السابع

المتغيرات العشوائية والتوزيعات الاحتمالية

1. متغيرات عشوائية منفصلة أو متقطعة
2. متغيرات عشوائية متصلة أو مستمرة

المتغير العشوائي:

ليكن فضاء العينة للتجربة العشوائية هو S.

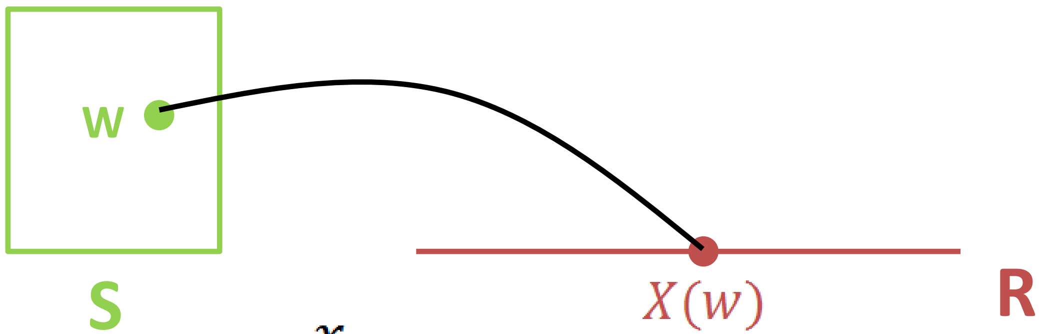
المتغير العشوائي X هو دالة حقيقية معرفة على فضاء العينة S.

ملاحظة:

1. لابد أن تتحقق بعض الشروط على الدالة لكي تكون متغيراً عشوائياً ولكننا لن نتطرق إلى تلك الشروط لأن ذلك خارج نطاق كتابنا هذا.
2. المتغير العشوائي X يعطي قيمة حقيقية وحيدة لكل عنصر من عناصر فضاء العينة S.
3. المتغير العشوائي X هو تطبيق مجاله فضاء العينة S ومجاله المقابل هو مجموعة الأعداد الحقيقية R، أي أن، $X:S \to R$.

إذا كانت $w \in S$ نقطة عينة فإن صورة w تحت تأثير المتغير العشوائي X هي $X(w)$ و هي قيمة حقيقية،

أي أن $X(w) \in R$ ، والشكل التالي يوضح صورة نقطة العينة w تحت تأثير المتغير العشوائي X.



$$w: \xrightarrow{x} X(w) \in R$$

إن المجموعة $X(S) = \{x \in R: X(w) = x, w \in S\}$ هي مدى التطبيق X و تسمى مجموعة القيم

الممكنة للمتغير العشوائي X، وهي مجموعة جزئية من مجموعة الأعداد الحقيقية أي أن $X(S) \subseteq R$.

مثال 1: لتكن التجربة هي قذف قطعة نقود متزنة مرتين متتاليتين بشكل مستقل. ولنعرف المتغير العشوائي X على أنه عدد الصور الظاهره في الرميتين.

1. عبر عن المتغير العشوائي X كدالة.

2. أوجد مجموعة القيم الممكنة للمتغير العشوائي X. (المدى)

3. عبر عن الحوادث التالية باستخدام المتغير العشوائي X:

{TT}, {HT, TH} , {HH}, {HH, HT, TH}

4. عبر عن الحوادث التالية باستخدام نقاط العينة:

$\{X=0\}, \{X=1\}$, $\{X=2\}, \{X<1\}, \{X<1\}, \{X>5\}$

5. أوجد الاحتمالات التالية:

$P(X=0)=$

$P(X=1)=$

$P(X=2)=$

5. أوجد الاحتمالات التالية:

$P(X<1)=$

$P(X\leq 1)=$

$P(X>5)=$

المتغير العشوائي المتقطع (المنفصل):

يكون المتغير العشوائي X متقطعاً إذا كانت مجموعة القيم الممكنة له مجموعة متقطعة.

ملاحظة:

مجموعة القيم الممكنة للمتغير العشوائي المتقطع X تاخذ إحدى حالتين:

1. $X(S) = \{x_1, x_2, \dots, x_n\}$
2. $X(S) = \{x_1, x_2, x_3, \dots\}$

مثال 2: في تجربة قذف قطعة نقود مرتين متتاليتين أوجد مجموعة القيم الممكنة للمتغيرات العشوائية التالية وحدد فيما إذا كانت متغيرات عشوائية متقطعة أم لا:

1. **المتغير العشوائي X الذي يمثل عدد الصور.**
2. **المتغير العشوائي Y الذي يمثل مربع عدد الصور.**
3. **المتغير العشوائي Z الذي يمثل عدد الصور مطروحاً منه عدد الكتابات.**

w	$X(w)$	$Y(w)$	$Z(w)$

مثال 3:في تجربة قذف قطعة نقود غير متزنة مرتين متتاليتين بشكل مستقل. ولنفرض أن هذه العملة غير متزنة بحيث أن $P(H)=1/3$ و $P(T)=2/3$. ولنعرف المتغير العشوائي X على أنه عدد الصور الظاهره في الرميتين.

1. أوجد الاحتمالات التالية ثم لخصها في جدول:

$P(X=0)=$

$P(X=1)=$

$P(X=2)=$

2. باستخدام (1) أوجد الاحتمالات التالية:

$P(0<X<2)=$

$P(X\leq 1)=$

$P(X\geq 2)=$

$P(X\geq 5)=$

$P(X<5)=$

دالة الكتلة الاحتمالية للمتغير العشوائي المتقطع:

إذا كان X متغيراً عشوائياً متقطعاً مجموعة القيم الممكنة له هي $X(S) = \{x_1, x_2, \dots, x_n\}$

أو $X(S) = \{x_1, x_2, x_3, \dots\}$ فإن دالة الكتلة الاحتمالية للمتغير العشوائي X يرمز لها بالرمز $f_X(x)$

أو $f(x)$ وتعرف كما يلي:

$$f_X(x) = \begin{cases} P(X = x); & x \in X(S) \\ 0; & x \notin X(S) \end{cases}$$

خواص دالة الكتلة الاحتمالية:

إن دالة الكتلة الاحتمالية $f_X(x)$ لابد أن تحقق الشروط التالية :

1. $0 \le f_X(x) \le 1$.

2. $\sum_{\forall x} f_X(x) = 1$

3. $P(X \in A) = \sum_{x \in A} f_X(x) = \sum_{x \in A} P(X \in A)\,;\ \forall\, A \subseteq R$

- **مثال 4: أوجد دالة الكتلة الاحتمالية للمتغير العشوائي X في مثال (3).**

X	$f_X(x) = P(X = x)$

التوقع (المتوسط) للمتغير العشوائي المتقطع:

إذا كان X متغيراً عشوائياً متقطعاً مجموعة القيم الممكنة له هي $X(S)=\{x_1,x_2,\dots,x_n\}$ أو $X(S)=\{x_1,x_2,x_3,\dots\}$ و دالة الكتلة الاحتمالية $f_X(x)$ فإن التوقع (أو القيمة المتوقعة أو المتوسط) للمتغير العشوائي X يرمز له بالرمز $E(X)$ أو بالرمز μ_x ويعرف بالصيغة التالية:

$$\mu_x = E(X) = \sum_{x\in X(S)} x f_X(x) = \sum_{x\in X(S)} xP(X=x)$$
$$= x_1 f(x_1) + x_2 f(x_2) + \cdots$$

مثال 5: أوجد توقع المتغير العشوائي X الذي دالة كتلته الاحتمالية معطاة في الجدول التالي:

X	$f_X(x)=P(X=x)$
0	4/9
1	4/9
2	1/9

بعض خواص التوقع:

ليكن X متغيراً عشوائياً ولتكن a و b ثوابت. إن التوقع يحقق الخواص التالية:

1. $E(a) = a$
2. $E(X \pm b) = E(X) \pm b$
3. $E(aX) = aE(X)$
4. $E(aX \pm b) = aE(X) \pm b$

نتيجة:

ليكن X متغيراً عشوائياً متقطعاً ولتكن $g(x)$ دالة حقيقية في المتغير العشوائي X. إن توقع الدالة يمكن حسابة بالصيغة التالية:

$$\mu_{g(x)} = E[g(x)] = \sum_{x \in X(S)} g(x) f_X(x) = g(x_1) f(x_1) + g(x_2) f(x_2) + \cdots$$

و كحالة خاصة عندما تكون $g(x) = X^2$ فإن:

$$E(X^2) = \sum_{x \in X(S)} x^2 f_X(x) = x_1^2 f(x_1) + x_2^2 f(x_2) + \cdots$$

مثال 6: أوجد توقع المتغيرات العشوائية التالية في مثال 5:

أ. $g(X) = 9X + 2$

ب. $g(X) = X^2$

x	$f_X(x) = P(X = x)$	$x^2 f_X(x)$
0	4/9	
1	4/9	
2	1/9	

التباين للمتغير العشوائي:

إذا كان X متغيراً عشوائياً توقعه (متوسطه) μ_x. فإن تباين المتغير العشوائي X يرمز له بالرمز V(X) أو Var(X) أو σ_x^2 ويعرف بالصيغة التالية:

$$\sigma_X^2 = V(X) = E[(x - \mu_x)^2]$$

الانحراف المعياري للمتغير العشوائي:

يرمز للانحراف المعياري للمتغير العشوائي X بالرمز σ_X ويعرف بالصيغة التالية:

$$\sigma_X = \sqrt{\sigma_X^2}$$

التباين للمتغير العشوائي المتقطع :

إذا كان X متغيراً عشوائياً متقطعاً دالة كتلته الاحتمالية هي $f_X(x)$ وتوقعه μ_x فإن تباين المتغير العشوائي X يمكن أن يحسب بالصيغة التالية:

$$\sigma_X^2 = Var(X) = \sum_{x \in X(S)} (x - \mu_x)^2 f_X(x) = \sum_{x \in X(S)} (x - \mu_x)^2 P(X = x)$$

صيغة حسابية للتباين:

باستخدام خواص التوقع فإنه يمكن إثبات أن التباين يعطى بالصيغة الحسابية التالية:

$$\sigma_X^2 = Var(X) = E(X^2) - [E(x)]^2 = E(X^2) - {\mu_x}^2$$

حيث أن: $E(X^2) = \sum_{x \in X(S)} x^2 f_X(x)$

مثال 7: أحسبي المتوسط والتباين والانحراف المعياري للمتغير العشوائي X الذي دالة كتلته الاحتمالية معطاة في الجدول أدناه:

X	$f_X(x)$	$x^2 f_X(x)$
0	0.6	
1	0.3	
2	0.1	
المجموع		

بعض خواص التباين:

ليكن X متغيراً عشوائياً ولتكن a و b ثوابت. إن التوقع يحقق الخواص التالية:

1. $Var(a) = 0$
2. $Var(X \pm b) = Var(X)$
3. $Var(aX) = a^2 Var(X)$
4. $Var(aX \pm b) = a^2 Var(X)$

مثال 8: أحسبي تباين المتغيرات العشوائية التالية في مثال (7).

أ. g(X)=10X

ب. g(X)=10X+2

بعض التوزيعات الاحتمالية المتقطعة :

التوزيعات الاحتمالية المتقطعة هي توزيعات احتمالية (أو دوال كتل احتمالية) لمتغيرات عشوائية متقطعة.

1. توزيع برنوللي
2. توزيع ذي الحدين
3. توزيع بواسون

محاولة برنوللي (Bernoulli's Trial):

تجربة عشوائية لها نتيجتين فقط.

نسمي الأولى بالنجاح ونرمز لها بالرمز s. ونرمز لاحتمال النجاح بالرمز $p=P(s)$

والثانية نسميها بالفشل ونرمز لها بالرمز f. ونرمز لاحتمال الفشل بالرمز $q=P(f)$

لذا فإن فراغ العينة لمحاولة برنوللي هو $S=\{s,f\}$.

ملاحظة:

$q=1-p$

من الأمثلة على محاولات برنوللي:

- رمي قطعة نقود (صورة أو كتابة)
- نتيجة الطالب في الاختبار (نجاح أو رسوب)
- فحص قطعة من إنتاج أحد المصانع (سليمة أو تالفة)

توزيع برنوللي:

لنفرض أن لدينا محاولة برنوللي و لنعرف المتغير العشوائي X على أنه عدد مرات النجاح عند إجراء محاولة برنوللي، أي أن :

$$X(s)=1, \quad X(f)=0$$

مجموعة القيم الممكنة للمتغير العشوائي X هي $X(S)=\{0,1\}$.

واحتمالاته هي:

$$P(X=1)=P(s)=p$$

$$P(X=0)=P(f)=q$$

بالتالي:

X	f(x)=P(X=x)
0	1-p
1	p

أي أن دالة الكتلة الاحتمالية للمتغير العشوائي X هي:

$$f_X(x) = P(X = x) = \begin{cases} p^x q^{1-x}; & x = 0,1 \\ 0; & x \neq 0,1 \end{cases}$$

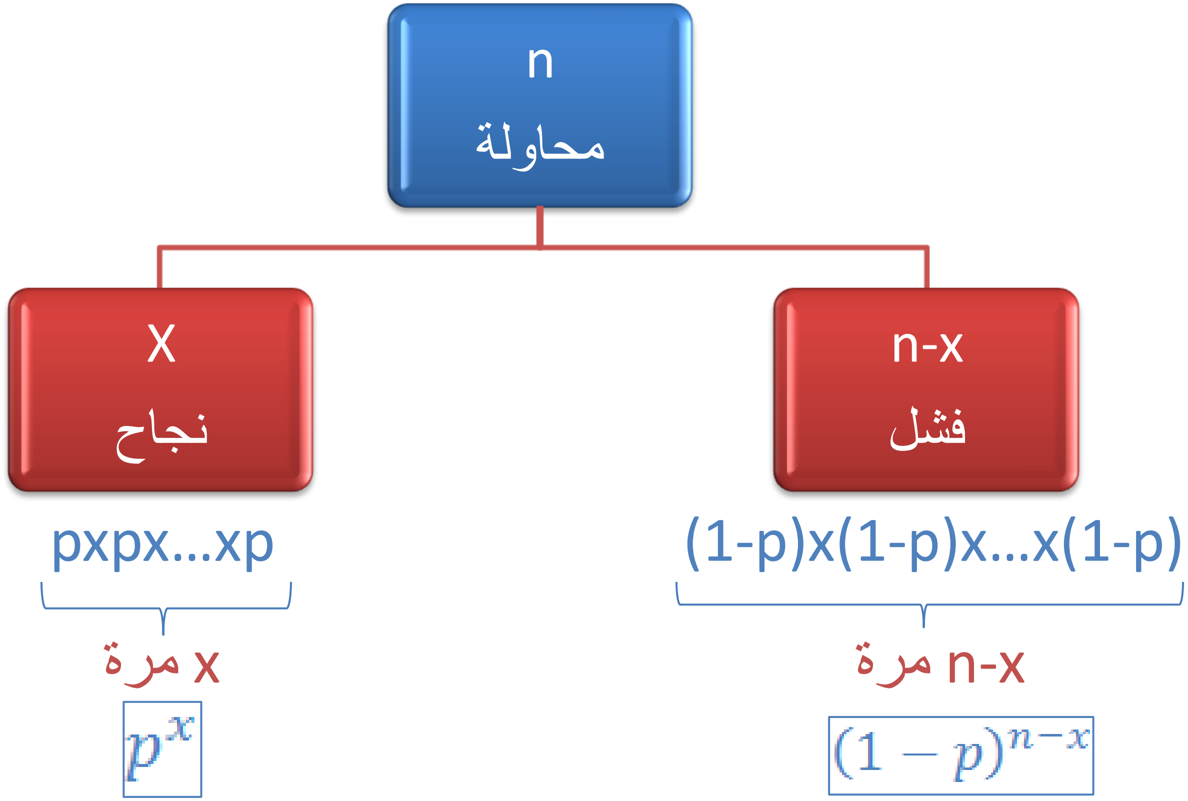
- معلمة هذا التوزيع هي احتمال النجاح p.

توزيع ذي الحدين:

لنفرض أن التجربة العشوائية تتكون من تكرار محاولة برنوللي عدد من المرات تحت الشروط التالية:

1. عدد المحاولات =n
2. المحاولات مستقلة
3. احتمال النجاح p ثابت لجميع المحاولات

لنعرف المتغير العشوائي X على أنه عدد مرات النجاح عند تكرار إجراء محاولة برنوللي. فتكون مجموعة القيم لهذا المتغير العشوائي هي X(S)={0,1,2,...,n}.



دالة الكتلة الاحتمالية للمتغير X هي:

$$f_X(x) = P(X = x) = \begin{cases} \binom{n}{x} p^x q^{n-x}; & x = 0,1,\dots,n. \\ 0; & x \neq 0,1,\dots,n. \end{cases}$$

- معلمتا هذا التوزيع هما عدد المحاولات n و احتمال النجاح p.

ملاحظة:

1. توزيع برنوللي هو حالة خاصة من توزيع ذي الحدين عندما n=1.
2. ليكن المتغير X هو عدد مرات النجاح وليكن المتغير Y هو عدد مرات الفشل، أي أن Y=n-X. إذا كان X يتوزع وفق توزيع Binomial(n,p)، فإن Y يتوزع وفق توزيع Binomial(n,1-p).

التوقع والتباين لتوزيع ذي الحدين:

إذا كان X متغيراً عشوائياً يتوزع وفق التوزيع ذي الحدين بالمعلمتين n و p فإن المتوسط والتباين للمتغير X هما على التوالي:

$$\mu_x = E(X) = \sum_{x=0}^{n} x f_X(x) = np$$

$$\sigma_X^2 = Var(X) = E(X^2) - {\mu_x}^2 = npq$$

مثال 9: لنفرض أن لدينا عملة غير متزنة بحيث أن $P(H)=0.4$ و $P(T)=0.6$. رميت هذه العملة ثلاث مرات بشكل مستقل. ليكن المتغير X يمثل عدد مرات ظهور الصورة في الرميات الثلاث.

أ. أوجدي دالة الكتلة الاحتمالية للمتغير العشوائي X.

ب. أوجدي التوقع (المتوسط) والتباين للمتغير العشوائي X.

ج. أوجدي الاحتمالات التالية:

1. الحصول على صورتين.

2. الحصول على صورتين على الأقل.

3. الحصول على صورة واحدة على الأكثر.

4. الحصول على ثلاث كتابات.

مثال 10: نسبة الإنتاج التالف لأحد مصانع المصابيح هي 10%. إذا أخذت عينة مكونة من 5 مصابيح بشكل عشوائي من إنتاج هذا المصنع، فأوجد مايلي:

أ. أوجدي الاحتمالات التالية:

1. الحصول على مصباح واحد تالف.

2. الحصول على جميع المصابيح تالفة.

3. الحصول على مصباح واحد تالف على الأكثر.

4. الحصول على مصباح واحد تالف على الأقل.

ب. أوجدي العدد المتوقع والتباين للمصابيح التالفة في العينة.

تجربة بواسون:

عملية ينتج عنها قيم عددية لمتغير عشوائي X يمثل عدد مرات وقوع حادثة ما في فترة زمنية (أو حيز مكاني) يرمز لها بالرمز t.

من الأمثلة على متغيرات بواسون:

- عدد الأخطاء المطبعية لكل صفحة في الكتاب. (t=صفحة واحدة)
- عدد المكالمات الهاتفية لسنترال إحدى الدوائر الحكومية لكل يوم. (t=يوم واحد)
- عدد المكالمات الهاتفية لسنترال إحدى الدوائر الحكومية لكل خمسة أيام. (t=خمسة أيام)

إذا فرضنا أن عدد مرات وقوع الحادثة في فترة الوحدة الزمنية (أو الحيز المكاني) (t=1) هو λ، فإن متوسط عدد مرات وقوع الحادثة في الفترة الزمنية (أو الحيز المكاني) $\mu = \lambda t$.

لنفرض أن المتغير X يمثل عدد مرات وقوع حادثة ما في الفترة الزمنية (أو الحيز المكاني) t.

توزيع بواسون:

يقال بأن المتغير العشوائي X يتوزع وفق توزيع بواسون بالمعلمة μ (حيث $\mu = \lambda t$) إذا كانت دالة كتلته الاحتمالية تعطى بالصورة:

$$f_X(x) = P(X = x) = \begin{cases} \dfrac{e^{-\mu}\mu^x}{x!}; & x = 0,1,2,\dots. \\ 0; & x \neq 0,1,2,\dots. \end{cases}$$

وفي هذه الحالة نكتب $X \sim Poisson(\mu)$.

- معلمة هذا التوزيع هو متوسط عدد مرات وقوع الحادثة μ.

التوقع والتباين لتوزيع بواسون:

إذا كان X متغيراً عشوائياً يتوزع وفق توزيع بواسون بالمعلمة μ، أي أن $X \sim Poisson(\mu)$

فإن:

1. $E(X) = \mu = \lambda t$
2. $Var(X) = \mu = \lambda t$

ملاحظة:

1. إن λ ماهو إلا متوسط توزيع بواسون للمتغير الذي يمثل وقوع الحادثة في فترة الوحدة.

2. لنفرض أن المتغير X يمثل عدد الأعاصير التي تضرب إحدى الجزر في الشهر (الوحدة الزمنية هي الشهر) و لنفرض أن المتغير Y يمثل عدد الأعاصير التي تضرب الجزيرة في السنة ولنفرض أن W يمثل عدد الأعاصير التي تضرب الجزيرة في اليوم. إذا كان المتغير X يتوزع وفق توزيع بواسون بالمعلمة λ، أي أن $X \sim Poisson(\lambda)$، فإن:

أ. المتغير العشوائي Y يتوزع وفق توزيع بواسون بالمعلمة $\mu = 12\lambda$، أي أن:

$$Y \sim Poisson(12\lambda) \quad .$$

ب. المتغير العشوائي W يتوزع وفق توزيع بواسون بالمعلمة $\mu = \frac{\lambda}{30}$، أي أن:

$$W \sim Poisson\left(\frac{\lambda}{30}\right)$$

- مثال 11: قام أحد المتدربين على الطباعة بطباعة كتاب. ولنفرض أن عدد الأخطاء المطبعية لكل صفحة في هذا الكتاب يتوزع وفق توزيع بواسون بمتوسط 6 أخطاء لكل صفحة.

1. إذا اخترنا أحد الصفحات من هذا الكتاب بشكل عشوائي، فما هو احتمال :

أ. أن يوجد 7 أخطاء مطبعية في هذه الصفحة؟

ب. أن يوجد مالا يقل عن خطأين مطبعيين في هذه الصفحة؟

2. إذا اخترنا صفحتين اثنتين من صفحات هذا الكتاب بشكل عشوائي، فما هو احتمال أن يوجد 10 اخطاء مطبعية فيهما؟

3. إذا اخترنا نصف صفحة من صفحات هذا الكتاب بشكل عشوائي، فما هو احتمال أن لا يوجد فيها أخطاء مطبعية؟

4. ماهو متوسط عدد الأخطاء المطبعية لكل سطر من أسطر هذا الكتاب بافتراض أن كل صفحة تحتوي 32 سطراً؟

5. ماهو توزيع ومتوسط و تباين المتغير العشوائي X الذي يمثل عدد الأخطاء المطبعية لكل 10 صفحات؟

دالة التوزيع التراكمية للمتغير العشوائي المتقطع (المنفصل):

إذا كان X متغيراً عشوائياً دالة كتلته الاحتمالية هي f(x)، فإن دالة توزيعه التراكمية يرمز لها بالرمز F(x) وتعرف لكل قيمة حقيقية x بالصيغة التالية:

$$F(x) = P(X \leq x) = \sum_{t \leq x} P(X = t) = \sum_{t \leq x} f(t) ; -\infty < x < \infty$$

أي أن قيمة دالة التوزيع التراكمية للمتغير العشوائي المتقطع X عند النقطه x تساوي احتمال أن يأخذ المتغير العشوائي X قيمة أصغر من أو يساوي القيمة x.

نتيجة:

إذا كان X متغيراً عشوائياً متقطعاً دالة توزيعه التراكمية هي F(x)، فإن:

$$P(a < X \leq b) = P(X \leq b) - P(X \leq a) = F(b) - F(a)$$

مثال 13: ليكن X متغيراً عشوائياً متقطعاً دالة كتلته الاحتمالية معطاة كما يلي:

X	0	1	2
f(x)	10/28	15/28	3/28

1. **أوجدي دالة التوزيع التراكمية للمتغير العشوائي X.**

2. **باستخدام دالة التوزيع التراكمية، أوجدي الاحتمال** $P(0.5 < X \leq 1.9)$.

المتغير العشوائي المستمر (المتصل):

هو متغير عشوائي مجموعة القيم الممكنة له عبارة عن فترة أو اتحاد عدد من الفترات. ومن أمثلة الكميات التي يمكن تمثيلها بواسطة متغيرات عشوائية متصلة:

- درجة حرارة تفاعل كيميائي معين
- نسبة تركيز مركب ما في محلول كيميائي
- الفترة الزمنية بين الإصابة بمرض ما والوفاة
- المسافة المقطوعة لجسم معين خلال وحدة الزمن
- طول الشخص

دالة الكثافة الاحتمالية للمتغير العشوائي المستمر:

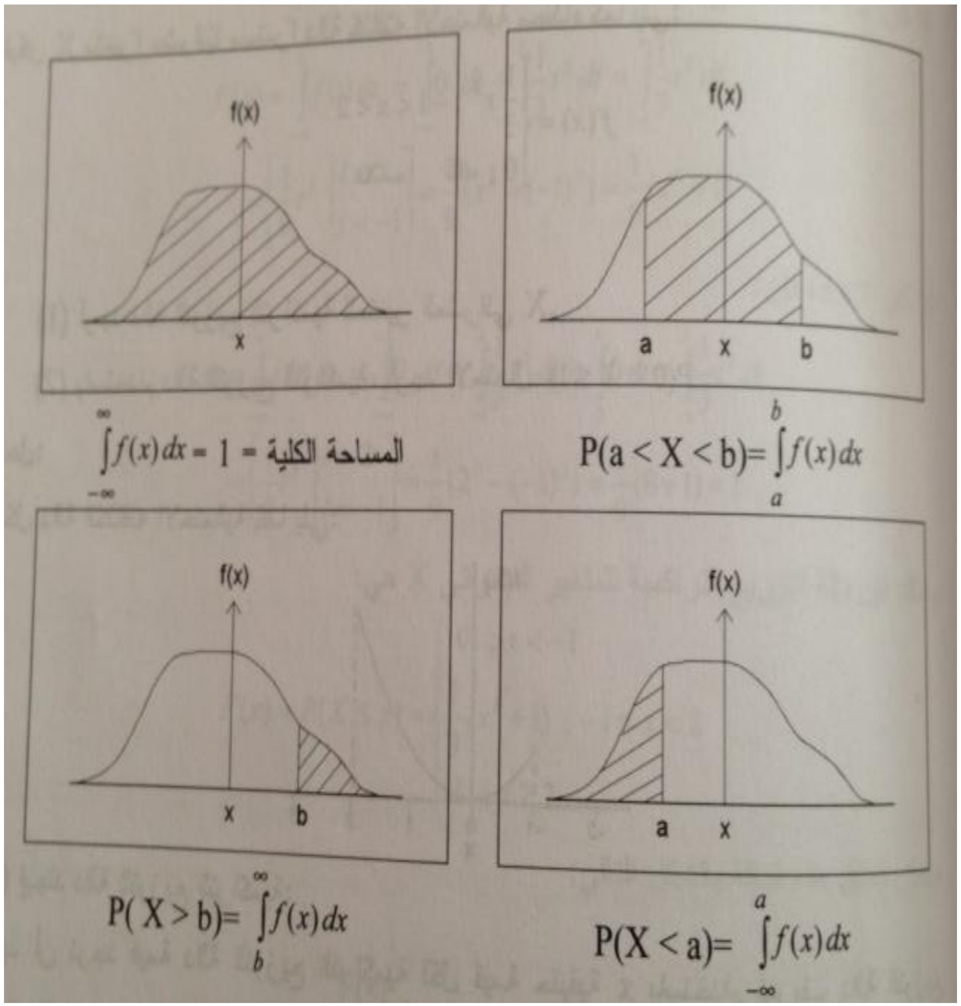
تعريف:

أي دالة حقيقية غير سالبة $f_X(x)$ والمعرفة على مجموعة الأعداد الحقيقية **R** تسمى دالة كثافة احتمالية إذا وفقط إذا كان:

$$P(a \leq X \leq b) = \int_a^b f_X(x)dx;\ \forall a, b \in R, a \leq b$$

أي أن احتمال وقوع المتغير العشوائي X في أي فترة يساوي المساحة فوق تلك الفترة وتحت منحنى الدالة.

ملاحظة: إذا كان X متغير عشوائي مستمرا دالة كثافته الاحتمالية $f_X(x)$ ، فإن:



- $f_X(x) \neq P(X = x)$
- $f_X(x) \geq 0 \ \forall x \in R$
- $P(X = x) = 0 \ \forall x \in R$
- $P(b < X \leq a) = P(b \leq X < a)$

$= P(b < X < a) = P(b \leq X \leq a)$

- $\int_{-\infty}^{\infty} f_X(x)dx = 1$

دالة التوزيع التراكمي للمتغير العشوائي المستمر:

تعريف:

إذا كان X متغير عشوائي مستمرا دالة كثافته الاحتمالية $f_X(x)$ ، فإن دالة توزيعه التراكمية ويرمز لهل بالرمز $F(x)$ وتعرف لكل قيمة حقيقة x

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f \quad (t)dt$$

ملاحظة:

$$P(a \leq X \leq b) = F(b) - F(a)$$

دالة التوزيع التراكمي للمتغير العشوائي المستمر:

مثال:

ليكن X متغير عشوائي مستمرا دالة كثافته الاحتمالية معطاه كما يلي

$$f(x) = \begin{cases} \frac{1}{3}x^2; -1 < x < 2 \\ 0 \text{ ; خلاف ذلك} \end{cases}$$

١. أوجد دالة التوزيع التراكمية للمتغير العشوائي X

٢. باستخدام دالة التوزيع التراكمية أوجد الاحتمال

$$P(0 < x \leq 1)$$

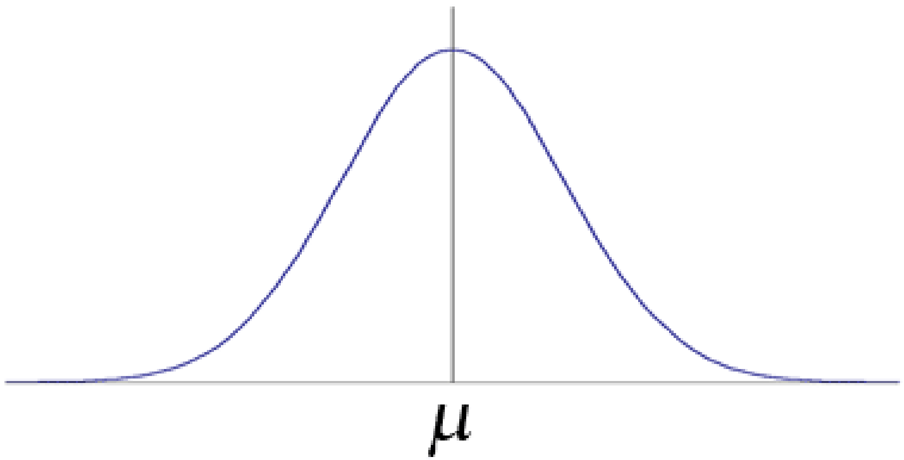
التوزيع الطبيعي (الاعتيادي):

يقال أن المتغير العشوائي المستمر X يتوزع وفق التوزيع الطبيعي بمتوسط μ وتباين σ^2 إذا كانت دالة كثافته الاحتمالية $f_X(x)$ وتأخذ الصيغة التالية:

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

حيث: $\sigma > 0 \qquad -\infty < \mu < \infty, \qquad -\infty < x < \infty,$

وفي هذه الحالة نكتب: $X \sim N(\mu, \sigma^2)$



دالة الكثافة الاحتمالية للتوزيع الطبيعي

ملاحظة:

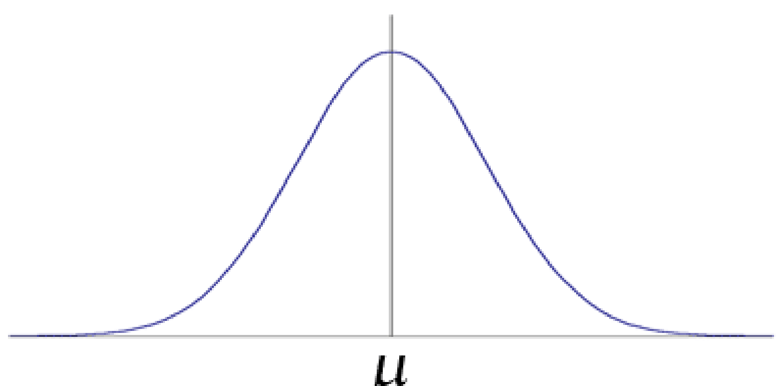
1. منحنى دالة الكثافة الاحتمالية للتوزيع الطبيعي $N(\mu,\sigma^2)$ متماثل حول المتوسط μ

2. نقطتا انقلاب دالة الكثافة الاحتمالية للتوزيع الطبيعي $N(\mu,\sigma^2)$ هما $x=\mu+\sigma$ و $x=\mu-\sigma$

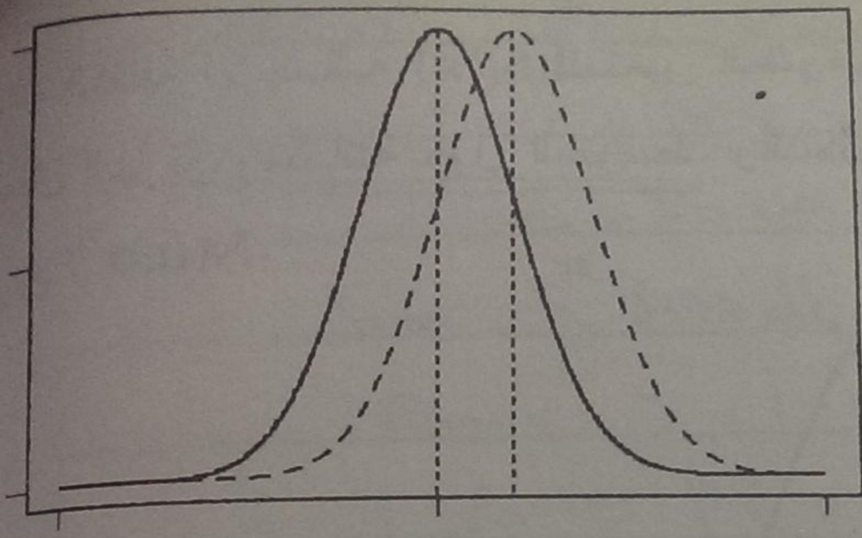
3. يتميز التوزيع الطبيعي $N(\mu,\sigma^2)$ بأن المتوسط = الوسيط = المنوال = μ

4. تعتمد دالة الكثافة الاحتمالية للتوزيع الطبيعي $N(\mu,\sigma^2)$ على معلمتي التوزيع و هما المتوسط μ والتباين σ^2 لذلك نكتب: $X \sim N(\mu,\sigma^2)$. و هاتان المعلمتان تحددان التوزيع الطبيعي تحديدا تاماً.
حيث أن المعلمة μ تحدد موضع التوزيع و المعلمة σ^2 تحدد شكل و تشتت التوزيع.

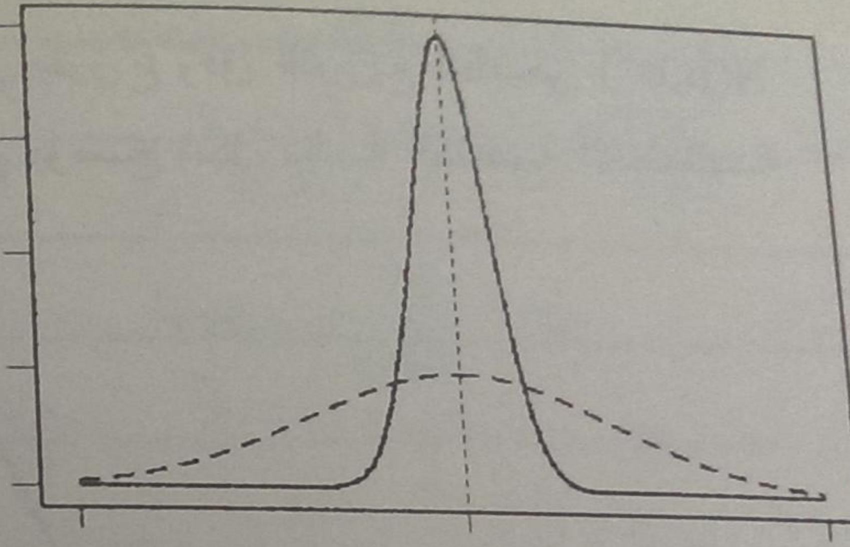
5. المساحة الكلية تحت منحنى دالة الكثافة الاحتمالية للتوزيع الطبيعي $N(\mu,\sigma^2)$ تساوي الواحد



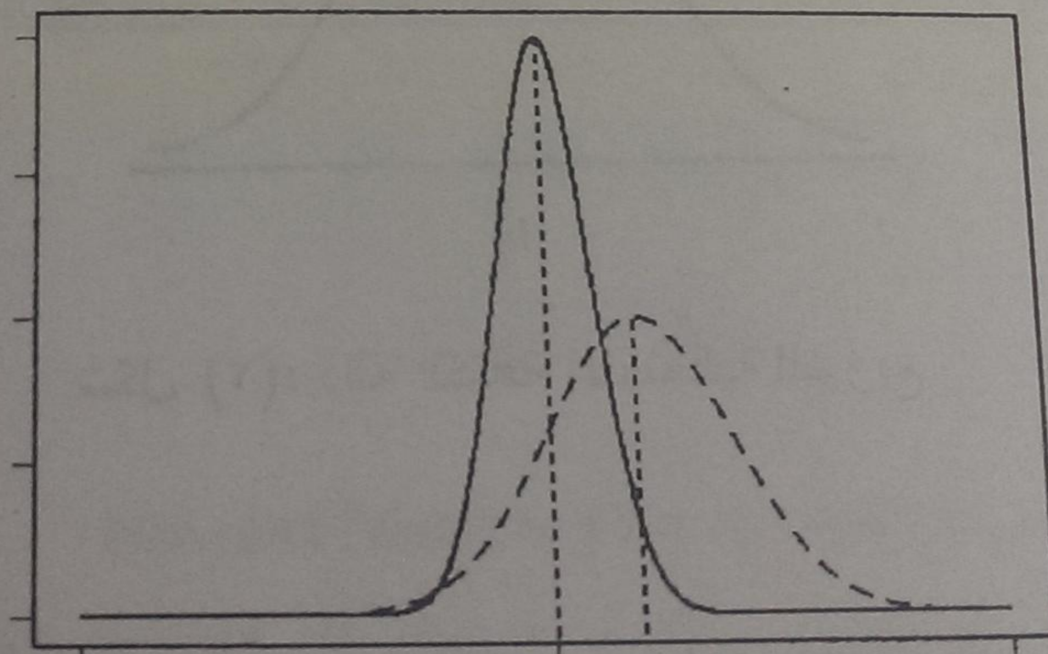
دالة الكثافة الاحتمالية للتوزيع الطبيعي



$\mu_1 < \mu_2, \sigma^2_1 = \sigma^2_2$

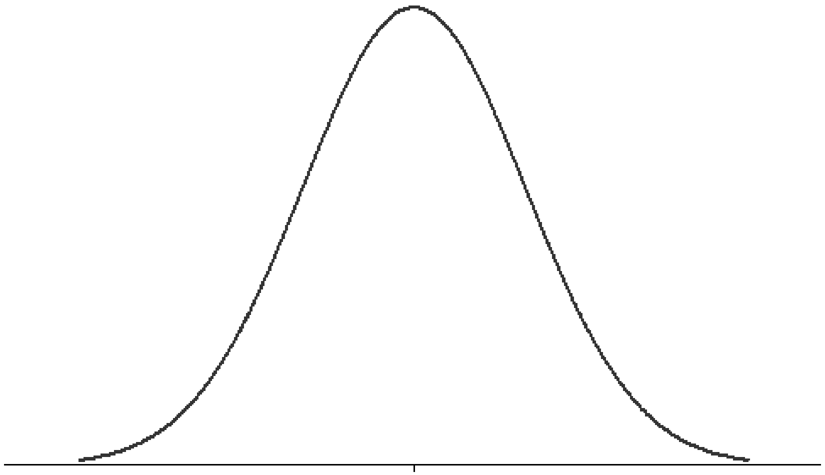


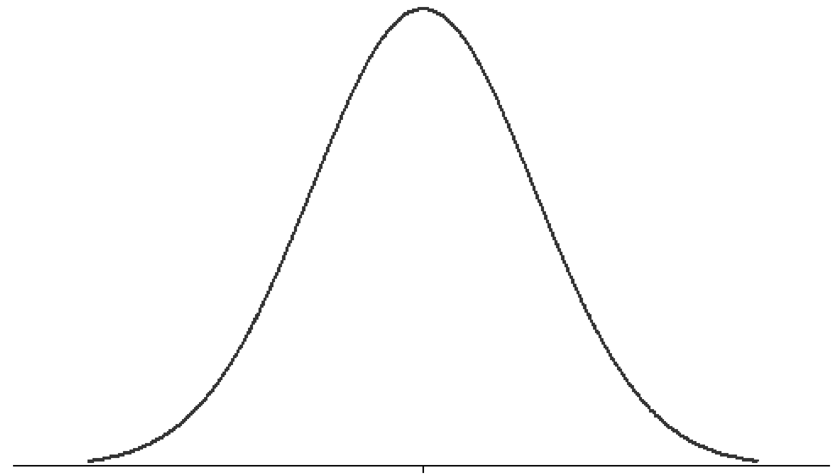
$\mu_1 = \mu_2, \sigma^2_1 < \sigma^2_2$

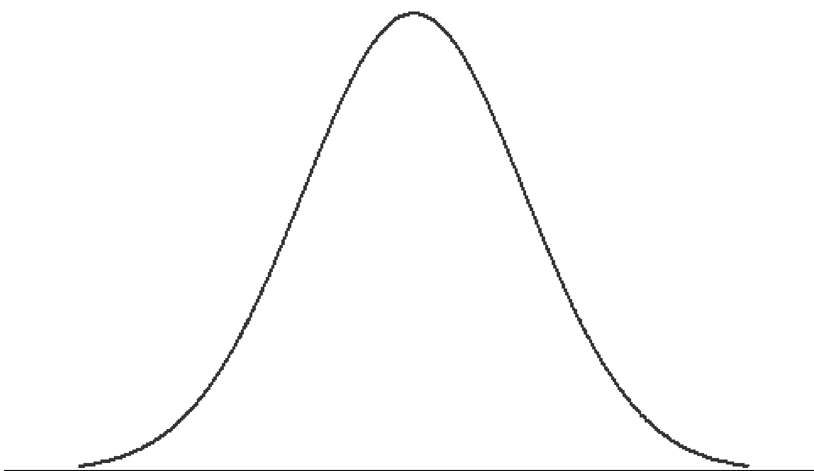


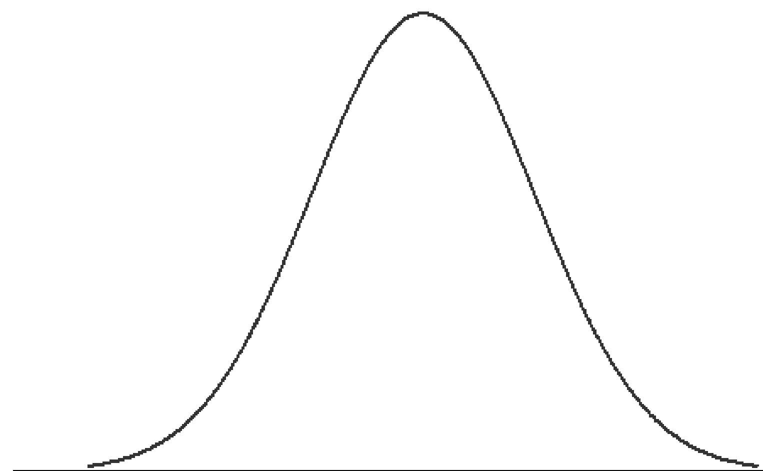
$\mu_1 < \mu_2, \sigma^2_1 < \sigma^2_2$

مثال 15: إذا كان طول الشخص في مجتمع ما X يتوزع تقريباً وفق التوزيع الطبيعي بمتوسط 160 سم وانحراف معياري 5 سم. مثل الاحتمالات التالية بمساحات تحت منحنى التوزيع الطبيعي:







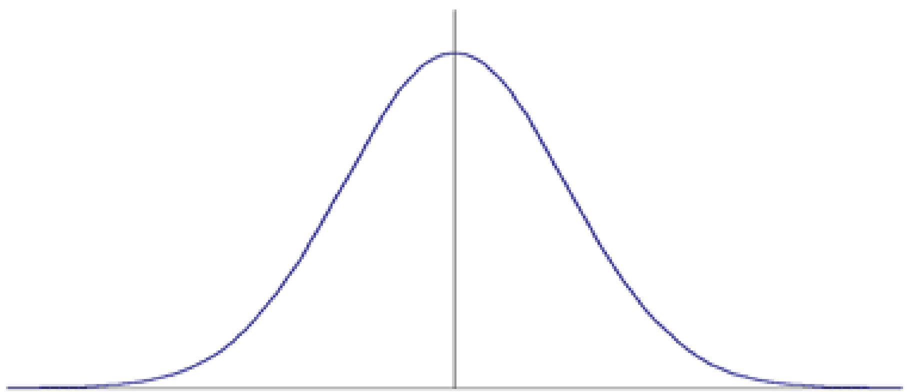


التوزيع الطبيعي المعياري (القياسي):

يقال أن المتغير العشوائي المستمر Z يتوزع وفق التوزيع الطبيعي المعياري إذا كان يتوزع وفق التوزيع الطبيعي بمتوسط يساوي الصفر ($\mu = 0$) وتباين يساوي الواحد ($\sigma^2 = 1$). ودالة الكثافة الاحتمالية للمتغير العشوائي Z تأخذ الصيغة التالية:

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\} \; ; \; -\infty < z < \infty$$

وفي هذه الحالة نكتب: $Z \sim N(0,1)$



دالة الكثافة الاحتمالية للتوزيع الطبيعي القياسي

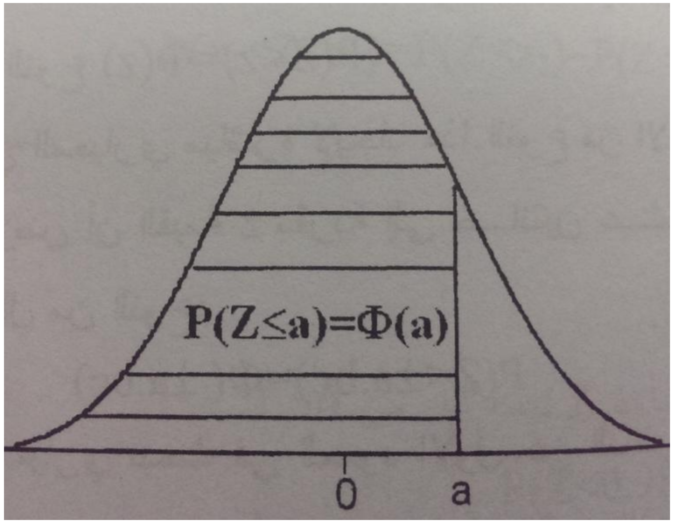
0

إيجاد الاحتمالات للتوزيع الطبيعي المعياري $Z \sim N(0,1)$:

إذا كان المتغير العشوائي Z يتوزع وفق التوزيع الطبيعي المعياري، فإن:

$$P(Z \le a) = \int_{-\infty}^{a} f_Z(z)dz = \int_{-\infty}^{a} \frac{1}{\sqrt{2\pi}} \exp\left\{-\frac{1}{2}z^2\right\} dz$$

وهذا التكامل يساوي المساحه المحصورة تحت منحنى الدالة $f_Z(z)$ وعن **يسار** النقطة a.



يرمز للاحتمال $P(Z \le a)$ بالرمز $\Phi(a)$ أي أن:

$$\Phi(a) = P(Z \le a) = \int_{-\infty}^{a} f_Z(z)dz$$

طريقة حساب الاحتمالات من جدول التوزيع الطبيعي القياسي:

1. $P(Z \leq z) = \Phi(z)$
2. $P(Z \geq z) = 1 - P(Z < z) = 1 - \Phi(z)$
3. $P(z_1 < Z < z_2) = P(Z < z_2) - P(Z < z_1) = \Phi(z_2) - \Phi(z_1)$
4. $P(Z < 0) = P(Z > 0) = \Phi(0) = 0.5$
5. $P(Z = z) = 0$

مثال 16: لنفترض أن $Z \sim N(0, 1)$.

1. أوجد احتمال أن يأخذ Z قيمة أقل من 1.5.
2. أوجد الاحتمال: $P(Z < 0.98)$
3. أوجد الاحتمال: $P(Z > 0.98)$
4. أوجد الاحتمال: $P(-1.33 < Z < 2.42)$

تحويل التوزيع الطبيعي $N(\mu, \sigma^2)$ إلى توزيع طبيعي معياري $N(0, 1)$ وإيجاد الاحتمالات للتوزيع الطبيعي $X \sim N(\mu, \sigma^2)$:

لإيجاد احتمالات المتغير العشوائي الطبعي $X \sim N(\mu, \sigma^2)$ فإننا نحوله أولا إلى متغير عشوائي طبيعي معياري $Z \sim N(0,1)$ ومن ثم نستخدم جدول التوزيع الطبيعي المعياري لإيجاد الاحتمالات من النوع $P(Z \leq z) = \Phi(z)$

نتيجة:

المتغير العشوائي X يتوزع وفق التوزيع الطبيعي $N(\mu, \sigma^2)$ إذا وفقط إذا كان المتغير العشوائي $Z = \dfrac{X - \mu}{\sigma}$ يتوزع وفق التوزيع الطبيعي المعياري $N(0,1)$. أي أن:

$$X \sim N(\mu, \sigma^2) \Leftrightarrow Z = \frac{X - \mu}{\sigma} \sim N(0,1)$$

فمثلاً: إذا كان $X \sim N(10,16)$ فإن

$$Z = \frac{X - 10}{4} \sim N(0,1)$$

ملاحظة:

1. باستخدام النتيجة السابقة فإن:

$$X < x \Leftrightarrow \frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma} \Leftrightarrow Z < \frac{x-\mu}{\sigma}$$

وبالتالي فإن:

1. $$P(X < x) = P\left(\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) = P\left(Z < \frac{x-\mu}{\sigma}\right) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$

2. $$P(X > x) = 1 - P(X < x) = 1 - P\left(\frac{X-\mu}{\sigma} < \frac{x-\mu}{\sigma}\right) = 1 - P\left(Z < \frac{x-\mu}{\sigma}\right) = 1 - \Phi\left(\frac{x-\mu}{\sigma}\right)$$

3. $$\begin{aligned} P(x_1 < X < x_2) &= P(X < x_2) - P(X < x_1) \\ &= P\left(\frac{X-\mu}{\sigma} < \frac{x_2-\mu}{\sigma}\right) - P\left(\frac{X-\mu}{\sigma} < \frac{x_1-\mu}{\sigma}\right) \\ &= P\left(Z < \frac{x_2-\mu}{\sigma}\right) - P\left(Z < \frac{x_1-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x_2-\mu}{\sigma}\right) - \Phi\left(\frac{x_1-\mu}{\sigma}\right) \end{aligned}$$

2. إذا كان $X \sim N(\mu, \sigma^2)$ وكانت x هي قيمة المتغير العشوائي X فإن القيمة $Z = \frac{X-\mu}{\sigma}$ تسمى القيمة المعيارية (أو القياسية) للقيمة x.

مثال 17: إذا كان المتغير العشوائي X الذي يمثل الطول في أحد المجتمعات البشرية يتوزع وفق التوزيع الطبيعي بمتوسط 165 سم وانحراف معياري 5 سم. فأوجد مايلي:

1. **القيمة المعيارية للقيمة $x=172$.**

2. **القيمة x إذا كانت القيمة المعيارية هي $z=-0.52$.**

مثال 18: لنفرض أن وزن الحيوان الذكر البالغ لإحدى سلالات الخراف يتوزع وفق التوزيع الطبيعي بمتوسط 16 كلجم وانحراف معياري 0.9 كلجم.

1. **إذا اخترنا أحد الخراف من هذه السلالة بشكل عشوائي فما هو احتمال أن يزيد وزنه عن 14 كلجم.**

2. **ماهي النسبة المئوية للخراف من هذه السلالة التي يزيد وزنها عن 14 كلجم.**

3. **ماهي النسبة المئوية للخراف من هذه السلالة التي تتراوح أوزنها من 14 إلى 18 كلجم.**