جامعة بغداد
كلية التربية للعلوم الصرفة / أبن الهيثم

Chapter
1 + 2 + 3

# الأحصاء

## قسم الحاسبات

## المرحلة الأولى

أستاذ المادة

**م. د. غدير جاسم محمد**

# الفصل الأول: مقدمة في الأحصاء

**علم الإحصاء:** هو ذلك العلم الذي يعمل على أستخدام الأسلوب العلمي في طرق جمع البيانات وتبويبها وعرضها وتحليلها بهدف الحصول منها على استنتاجات وقرارات مناسبة.

**بعض مفاهيم الإحصاء**

I. **البيانات (Data):** هي سلسلة غير مترابطة من الحقائق الموضوعية، التي يمكن الحصول عليها عن طريق الملاحظة، أو عن طريق البحث والتسجيل.

**تقسم البيانات الى قسمين:-**

a) بيانات الغير المبوبة ( Ungrouped Data) : هي البيانات الأولية التي جمعت ولم تبوب.

b) بيانات المبوبة ( Grouped Data ) : هي البيانات التي نظمت في جدول توزيع تكراري.

---

II. **المتغيرات (Variables):** المتغير في علم الإحصاء هو الخاصية أو السمة التي تأخذ قيما أو مستويات مختلفة من فرد إلى آخر ويرمز للمتغير بالرمز "X".

**تقسم المتغيرات الى قسمين: -**

a) **المتغيرات الوصفية (Categorical Variables):** وهي المتغيرات التي يمكن قياسها بوسائط قياس مألوفة او هي تلك الظواهر التي لا يمكن قياسها مباشرة بالأرقام العددية مثل الحالة الاجتماعية (غني، فقير).

b) **المتغيرات الكمية (Numerical Variables) :** هي تلك الظواهر او الصفات التي يمكن قياسها مباشرة بالأرقام العددية مثل صفة الطول، الوزن والعمر.

**يمكن تصنيف المتغيرات الكمية الى صنفين:**

a) **المتغيرات المستمرة (Continuous Variables):** هي المتغيرات التي تأخذ المتغيرات فيها أي قيمة رقمية ضمن مدى معين. لو افترضنا ان اطوال مجموعة من الأشخاص تتراوح بين 130cm و 170cm ، فأننا نستطيع كتابة المتغير بصيغة:

$$130 \leq X \leq 170$$

b) **المتغيرات المتقطعة (المنفصلة)(Discrete Variables):** هي المتغيرات التي تأخذ المفردة فيها قيما متباعدة. لو فرضنا ان عدد افراد الاسرة في أربعة عوائل هي {٥،٤،٣،٢} فأننا نستطيع كتابة المتغير بصيغة:

X = 2, 3 ,4, 5

---

**III. <u>المجتمع (Population)</u>:-** هو جميع القيم او المفردات التي يمكن ان يأخذها المتغير.

**يتكون المجتمع على نوعان:**

a) **مجتمع محدد**: هو المجتمع الذي يمكن حصر عدد مفرداته. على سبيل المثال عدد الوحدات المنتجة في مصنع معين.

b) **مجتمع غير محدد:** هو المجتمع الذي لا يمكن حصر عدد مفرداته. على سبيل المثال عدد الطيور المهاجرة.

---

**IV. <u>العينة (Sample)</u>**: هي مجموعة من المفردات يتم اختيارها بطريقة معينة من المجتمع.

**<u>المجموع ( Summation)</u>**

اذا كانت { X1, X2, … , Xn} تمثل مفردات المتغير X فأن مجموع هذه المفردات يعبر عنها بالرمز:

$$\sum_{i=1}^{n} X_i = X_1 + X_2 + \cdots + X_n$$

**Example: Let X $= 6, 7, 4, 3$. Find $\sum X$ , $\sum X^2$ and $(\sum X)^2$**

<u>Solution</u>:-

$$\sum X = 6 + 7 + 4 + 3 = 20$$

$$\sum X^2 = 36 + 49 + 16 + 9 = 110$$

$$(\sum X)^2 = (20)^2 = 400$$

---

## العرض الجدولي (Tabular Presentation):

**جدول التوزيع التكراري (Frequency Distribution)**

1. **الجدول البسيط:** هو جدول بسيط يتكون من عمودين في العمود الأول تقسم فيه قيم المتغير الى مجموعات تدعى الفئات ( Classes) والعمود الثاني يبين فيه عدد مفردات كل فئة يسمى التكرارات (Frequencies) ولكل فئة حدان هما الحد الأعلى (Maximum Value) و الحد الأدنى(Minimum Value).

**مثال: الجدول التالي يمثل توزيع عدد من الطلبة في جامعة ما حسب اوزانهم.**

| Classes | Frequencies (Fi) |
|---|---|
| 60-62 | 85 |
| 63-65 | 15 |
| 66-68 | 45 |
| 69-71 | 27 |
| 72-74 | 8 |
| summation | 180 |

2. **الجدول المزدوج: -**

وهو الجدول الذي توزع فيه البيانات حسب صفتين او ظاهرتين او أكثر في نفس الوقت. فمثلا" الجدول المزدوج (لصفتين) يتآلف من الصفوف التي تمثل فئات او مجاميع احدى الصفتين والاعمدة التي تمثل فئات او مجاميع الصفة الأخرى اما المربعات التي تقابل الصفوف والاعمدة تحتوي على عدد المفردات او التكرارات المشتركة في فئات ومجاميع كلا الصفتين.

**مثال: الجدول الاتي يمثل توزيع عدد من طلبة كلية ما حسب صفتي الطول والوزن.**

| **Length \ weight** | **51-60** | **61-70** | **71-80** | Summation |
|---|---|---|---|---|
| **121-140** | 20 | 6 | 4 | 30 |
| **141-160** | 2 | 40 | 10 | 52 |
| **161-180** | 2 | 6 | 10 | 18 |
| Summation | 24 | 52 | 24 | 100 |

---

**طول الفئة (Class Length):** هو مقدار المدى بين حدي الفئة. وهي عبارة عن الفرق بين الحدين الأعلى (او الحدين الأدنى) للفئتين المتتاليتين.

**Class Length = Maximum value– Minimum value**

**مركز الفئة (Class Mid - Point):** وهو عبارة عن منتصف المدى بين حدي الفئة.

**Class -Mid-point = (Maximum value + Minimum value) / 2**

---

**التكرار النسبي (Relative Frequency)** : و هي الأهمية النسبة لكل فئة.

$$\textbf{التكرار النسبي لأي فئة} = \frac{\text{تكرار الفئة}}{\text{مجموع التكرارات}}$$

$$\textbf{Relative Frequency} = \frac{\textbf{Fi}}{\sum \textbf{Fi}}$$

---

**ملاحظة: التكرار المئوي = (التكرار النسبي) * 100**

**Percentage Frequency = Relative Frequency *100**

**المدرج التكراري (Histogram):**

عبارة عن مستطيلات راسية تمتد قواعدها على المحور الافقي لتمثل أطول الفئات بينما ارتفاعاتها تمثل تكرارات الفئات ولرسم مدرج تكراري تتبع الخطوات الاتية: -

a) نرسم المحور الافقي والمحور العمودي

b) نقسم المحور الافقي الى اقسام متساوية بمقياس رسم مناسب بحيث يشمل جميع الحدود الحقيقية للفئات بينما نقسم المحور العمودي الى أقسام متساوية بحيث تشمل أكبر التكرارات.

c) نرسم على كل فئة مستطيل راسي تمثل قاعدته طول تلك الفئة وارتفاعه يمثل تكرار تلك الفئة.

---

**المضلع التكراري (Frequency Polygon)**

عبارة عن خطوط مستقيمة متكسرة تصل بين النقاط. كل نقطة واقعة فوق مركز فئة على ارتفاع يمثل تكرار تلك الفئة.

**ملاحظة: لرسم المضلع التكراري يجب رسم المدرج التكراري أولا ثم تصنف القواعد العليا للمستطيلات والتي تمثل مراكز الفئات بنقاط ثم نصل هذه النقاط بمستقيمات.**

---

**المنحنى التكراري (Frequency Carve):**

لا تختلف فكرة رسم المنحني التكراري (Frequency Carve) عن المضلع التكراري (Frequency Polygon ) من حيث الأسلوب لكن الفرق الوحيد ما بينهما هو انه بدلا من توصيل النقاط (مركز الفئة، التكرار) بمستقيمات فانه يتم تمرير منحني بين النقاط.

---

**التمثيل البياني لجدول التوزيع التكراري التجميعي:**

وهو عبارة عن خطوط مستقيمة متكسرة تصل بين نقاط واقعة فوق الحدود الحقيقية للفئات واعلى ارتفاع يمثل التكرار التجميعي وهنالك نوعان من المضلع التكراري التجميعي:

**1- المضلع التكراري التجميعي التصاعدي:**

لرسم المضلع التكراري التجميعي التصاعدي نتبع الخطوات الاتية:

a) نرسم المحور الافقي والمحور العمودي

b) نقسم المحور الافقي الى اقسام متساوية بمقياس رسم مناسب بحيث يشمل جميع الحدود الحقيقية للفئات بينما نقسم المحور العمودي الى أقسام متساوية بحيث تشمل أكبر التكرارات التجمعية التصاعدية.

c) نحدد ثم نصل النقاط (مركز الفئة، التكرار التجميعي التصاعدي) بخطوط مستقيمة.

**2- المضلع التكراري التجميعي التنازلي:**

ويرسم المضلع التكراري التجميعي التنازلي بنفس الطريقة السابقة غير ان هذا المضلع يبدأ من اعلى قيمة (مجموع التكرارات) وينتهي بأقل قيمة.

**Example: The following table represents the number of 100 students at the Mathematics department with their weight:**

| Classes | Frequencies |
|---|---|
| 60-62 | 5 |
| 63-65 | 18 |
| 66-68 | 42 |
| 69-71 | 27 |
| 72-74 | 8 |
| المجموع | 100 |

Find:

1. Class length
2. Class Mid-point.
3. Class boundary.
4. Draw the Histogram and Frequencies Polygon.
5. Draw the increasing cumulative Frequencies Polygon.
6. Draw the decreasing cumulative Frequencies Polygon.

Solution:

| Classes | Frequencies | Relative Frequencies | Percentage Frequencies | Increasing cumulative Frequencies | Decreasing cumulative Frequencies |
|---|---|---|---|---|---|
| 60-62 | 5 | 0.05 | 5 | 5 | 100 |
| 63-65 | 18 | 0.18 | 18 | 23 | 95 |
| 66-68 | 42 | 0.42 | 42 | 65 | 77 |
| 69-71 | 27 | 0.27 | 27 | 92 | 35 |
| 72-74 | 8 | 0.08 | 8 | 100 | 8 |
| المجموع | | | | | |

# Problem Set (1):

**Q1)** In the following table, the classes represent the grads and the frequencies represent the number of students in statistics class.

| **Classes** | 1-20 | 21-40 | 41-60 | 61-80 | 81-100 |
|---|---|---|---|---|---|
| **Frequencies** | 3 | 8 | 15 | 9 | 2 |

Find:

1. Class length
2. Class Mid-point.
3. Class boundary.
4. Draw the Histogram and Frequencies Polygon.
5. Draw the increasing cumulative Frequencies Polygon.
6. Draw the decreasing cumulative Frequencies Polygon.

**Q2)** Choose five students from your classmates to create a data set that has three numerical variables (2 continuous variables and 1 discrete variable), and two categorical variables.

**Q3)** The following table represents the number of people who died in the last 10 years in a small city.

| **Classes: Ages** | 30-40 | 41-50 | 51-60 | 61-70 | 71-80 | 81-90 |
|---|---|---|---|---|---|---|
| **Frequencies** | 4 | 8 | 13 | 20 | 39 | 32 |

Find:

1. Draw the Histogram and Frequencies Polygon.
2. Draw the increasing cumulative Frequencies Polygon.
3. Draw the decreasing cumulative Frequencies Polygon.

### مقاييس النزعة المركزية:

1 ) **الوسط الحسابي (Mean)**: وهي القيمة الناتجة من قسمة مجموع القيم على عددها.

أ) **البيانات الغير المبوبة**: اذا كان لدينا n من القيم (X1, X2, … , Xn) فأن الوسط الحسابي لها هو:

$$\mathbf{Mean} = \bar{\mathbf{X}} = \frac{\sum \mathbf{X_i}}{\mathbf{n}}$$

ب) **البيانات المبوبة**: اذا كان لدينا n من القيم (X1, X2, … , Xn) التي تمثل مراكز الفئات في جدول التوزيع التكراري مع تكراراتها (F1, F2, … , Fn) على التوالي فأن الوسط الحسابي لها هو:

$$\mathbf{Mean} = \bar{\mathbf{X}} = \frac{\sum \mathbf{F_i X_i}}{\sum \mathbf{F_i}}$$

**Example: Find the mean for the following data?**

| Classes | Fi | Xi | Fi*Xi |
|---|---|---|---|
| **31 - 40** | **1** | **35.5** | 35.5 |
| **41 - 50** | **2** | **45.5** | 91.0 |
| **51 - 60** | **5** | **55.5** | 277.5 |
| **61 - 70** | **15** | **65.5** | 982.5 |
| **71 - 80** | **25** | **75.5** | 1887.5 |
| **81 - 90** | **20** | **85.5** | 1710.5 |
| **91 - 100** | **12** | **95.5** | 1146.5 |

**Solution: -**

$$\mathrm{Mean} = \bar{\mathrm{X}} = \frac{\sum \mathrm{F_i X_i}}{\sum \mathrm{F_i}} = \frac{6130}{80} = 76.62$$

2 ) **<u>الوسيط (Median):</u>** في البيانات الغير المبوبة اذا كان لدينا n من القيم (X1, X2,…, Xn) ، نرتب القيم ترتيبا" تصاعديا" او تنازليا"

i. إذا كان n عددا" فرديا" فان الوسيط هو القيمة التي موقعها هو $\frac{(n+1)}{2}$.

ii. اذا كان n عددا" زوجيا" فان الوسيط هو متوسط القيمتان التي موقعهما هو $\frac{n}{2}$ و $1+\frac{n}{2}$.

**Example1 : Find the median for the following data points?**

80, 82,76,87,84

**Solution: -**

نرتب القيم ترتيبا" تصاعديا:

76, 80, **82**,84,87

عددا" فريديا" n=5 , فأن الوسيط هو القيمة التي موقعها

$$\frac{(n+1)}{2} = \frac{(5+1)}{2} = 3$$

$$So, \quad Median = 82$$

**Example 2 : Find the median for the following data points?**

5, 4, 8, 7, 3, 12, 9, 2

**Solution: -**

نرتب القيم ترتيبا" تصاعديا:

2, 3, 4, **5, 7**, 8, 9, 12

n = 8 وهو عددا" زوجيا"، فان الوسيط هو متوسط القيمتان التي موقعهما هو $\frac{n}{2}$ و $1+\frac{n}{2}$

$1+\frac{n}{2} = 1+\frac{8}{2} = 5$ , $\frac{n}{2} = \frac{8}{2} = 4$

$$So, \quad Median = \frac{5+7}{2} = 6$$

3 ) **<u>المنوال (Mode)</u> :**

في البيانات الغير المبوبة اذا كان لدينا n من القيم (X1,X2,…,Xn)، فأن المنوال هي القيمة الأكثر تكرارا".

**Example: Find the mode for the following data points?**

5, 4, 8, 7, 3, 12, 9, 3, 4, 3

**Solution: -** Mode = 3

## **مقاييس التشتت او الاختلاف:**

ونقصد به التباعد او التقارب الموجود بين قيم المشاهدات لمتغير ما، وأيضا هي مقياس مدى تشتت قيم المشاهدات عن وسطها فكلما كان التشتت كبير دل ذلك على عدم التجانس بين القيم والعكس صحيح.

i. **المدى(Range) :** في البيانات الغير مبوبة المدى لمجموعة من الأرقام هو الفرق بين اعلى قيمة واقل قيمة من تلك المجموعة.

**Example: Find the Range for the following data points?**

9, 4, 7, 8, 16, 10, 20, 2

**Solution: -** Range = 20 – 2 = 18

---

ii. **الانحراف المعياري(Standard Deviation) :** في البيانات الغير مبوبة يمكن ايجاد الانحراف المعياري باستخدام القانون التالي:

$$\textbf{Standard Deviation} = \textbf{s.d.} = \frac{\sum|\mathbf{Xi} - \bar{\mathbf{X}}|}{\mathbf{n}}$$

---

iii. **التباين(Variance) :** هو مقياس للتشتيت الاحصائي للقيم الممكنة حول القيم المتوقعة، وهو مساوي لتربيع الانحراف المعياري.

$$\textbf{Variance} = (\textbf{Standard Deviation})^{\mathbf{2}}$$

**Example: Find the standard deviation and the variance for the following data points?**

9, 8, 6, 5, 7

**Solution: -** First we find the Mean

$$Mean = \bar{X} = \frac{\sum X_i}{n} = \frac{9+8+6+5+7}{5} = 7$$

Then,

$$\text{Standard Deviation} = \text{s.d.} = \frac{\sum|Xi - \bar{X}|}{n}$$

| Xi | Xi - $\bar{X}$ | \|Xi - $\bar{X}$\| |
|---|---|---|
| 9 | 2 | 2 |
| 8 | 1 | 1 |
| 6 | -1 | 1 |
| 5 | -2 | 2 |
| 7 | 0 | 0 |

Hence, Standard Deviation = $\mathrm{s.d.} = \frac{\sum|\mathrm{Xi}-\bar{\mathrm{X}}|}{\mathrm{n}} = \frac{2+1+1+2+0}{5} = \frac{6}{5}$

$$\mathrm{Variance} = \mathrm{Var} = (\frac{6}{5})^2 = \frac{36}{25}$$

---

**معامل الاختلاف (Coefficient of Variation):**

هو مقياس لتشتت أو تبعثر توزيع الاحتمال او التوزيع التكراري، ويعرف معامل الاختلاف باستخدام القانون التالي:

$$Coefficient\ of\ Variation = C.V. = \frac{s.d.}{\bar{X}} * 100$$

**Example:** In Math and Stat final exams, we found that for Math exam ( $\bar{X} = 78$ *and* $s.d. = 8$), and for Stat exam ( $\bar{X} = 73$ *and* $s.d. = 7.6$), in which subject the variation (C.V.) is greater?

**Solution: -**

For Math exam: $C.V. = \frac{s.d.}{\bar{X}} * 100 = \frac{8}{78} * 100 = 10.25\%$

For Stat exam: $C.V. = \frac{s.d.}{\bar{X}} * 100 = \frac{7.6}{73} * 100 = 10.41\%$

From above, we can see that the variation for Stat exam is greater.

---

**الدرجة القياسية (Z-Score):**

في الكثير من الأحيان نحتاج الى المقارنة بين مفردتين من مجموعتين مختلفتين وفي هذه الحالة يجب تحويل وحدات كل مفردة الى وحدات قياسية حتى تكون المقارنة ذات معنى، ويكون هذا باستخدام الوسط الحسابي والانحراف المعياري لكل مجموعة وتسمى هذه القيمة بدرجة القياس ويرمز لها بالرمز (Zi) حيث ان:

$$\mathbf{Zi = \frac{Xi - \bar{X}}{s.d.}}$$

**Example:** In a classroom, a student got 84 in Mathematic exam where the mean for all students was 76 with s.d. = 10, and the same student got 90 in Biology exam where the mean for all the students was 82 with s.d. = 16. In which subject , did the student perform better?

**Solution: -**

For Math exam:

$$Zi = \frac{Xi - \bar{X}}{s.d.} = \frac{84 - 76}{10} = 0.8$$

For Bio exam:

$$Zi = \frac{Xi - \bar{X}}{s.d.} = \frac{90 - 82}{16} = 0.5$$

The student performs better in Math exam (because Z-score is higher) even though his grade in Bio exam is greater.

---

**Definition:** The range of a data set is the difference between the maximum value and the minimum value of the data set.

**Definitions:** **Q1** is the first quartile of the data, **Q2 (Median)** is the second quartile of the data, **Q3** is the third quartile of the data.

**Definition:** Inter Quartile Range (**IQR**) represents the middle 50% of the data.

**(i.e., IQR = Q3 – Q1)**

## Five-Number Summary:

The Five-Number summary for any data set are:

**Minimum value, Q1, Q2 (Median), Q3, Maximum value.**

**Example:** Find the five-Number summary for the following data set then compute the IQR?

11, 74, 52, 99, 21, 38, 14, 26, 35, 56, 83, 91

**Solution: -** First we arrange the data points in increasing order

11, 14, 21, 26, 35, 52, 56, 74, 83, 92, 99

Minimum Value = 11

Q1 = 21

Q2 = 52

Q3 = 83

Maximum Value = 99

**The Five-Number summary are:** 11, 21, 52, 83, 99

**IQR = Q3 – Q1** = 83 -21 = 62

---

## Maximum and Minimum Usual Values:

Minimum Usual Value = Mean - 2 * s.d.

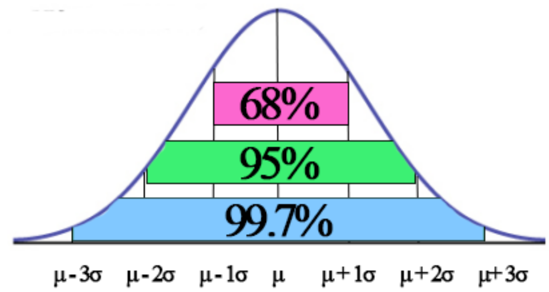
Maximum Usual Value = Mean + 2 * s.d.

**Example: Find the Maximum and Minimum usual values for a data set if you know the mean = 16 and the standard deviation = 3 ?**

Solution: -

Minimum Usual Value = Mean - 2 * s.d. = 16 – 2*3 = 10

Maximum Usual Value = Mean + 2 * s.d. = 16 + 2*3 = 22




**<u>Empirical Rule:</u>**

68 % of the data points with 1 s.d.

95 % of the data points with 2 s.d.

99.7 % of the data points with 3 s.d.

---

**Definition:** A value that lies between (Mean – 2*s.d.) and (Mean + 2*s.d.) is called a **<u>usual value</u>**, and any value lies outside this range is called **<u>unusual value</u>**.

---

**Example:** The mean of H2O bill for a house is \$40 with s.d. of \$4.34.

a) Find the maximum and minimum usual values?

b) Is an H2O bill of \$62 usual?

**Solution (a):**

Minimum Usual Value = Mean - 2 * s.d. = \$40 – 2*4.36 = \$48.72

Maximum Usual Value = Mean + 2 * s.d. = \$40 + 2*4.36 = \$31.28

**Solution (b):** H2O bill of \$62 is Unusual (Not usual) because it lies outside the usual range

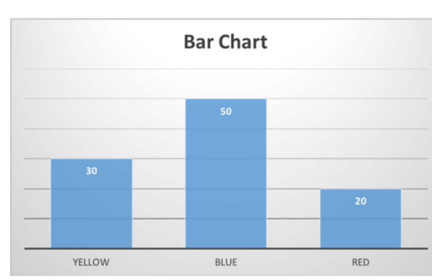
---

**<u>Data Visualization:</u>**

1. **Bar Chart:** A bar chart is a chart that presents categorical data with rectangular bars with heights proportional to the values that they represent.

**Example:** In a box, there are 100 balls such that:

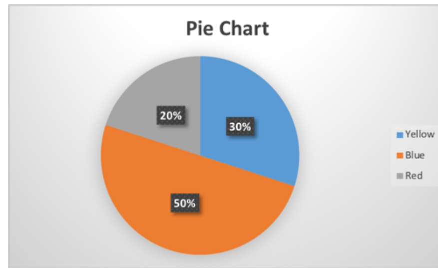
| Color | Frequency |
|---|---|
| Blue | 30 |
| Orange | 50 |
| Gray | 20 |



2. **Pie Chart:** A pie chart (or a circle chart) is a circular statistical graphic, which is divided into slices to illustrate numerical proportion.

**Example:** In a box, there are 100 balls such that:

| Color | Frequency |
|---|---|
| Blue | 30 |
| Orange | 50 |
| Gray | 20 |




3. **Box plot:** In descriptive statistics, a box plot or boxplot is a method for graphically depicting groups of numerical data through their quartiles.

**Example:** Create a box plot for the following. data points:

$$\{13,15,1,4,7,5,5,9,13,21,15\}$$

<u>**Solution:**</u>

First, we arrange the data points in increasing order:

**1, 4, 5, 5, 7, 9, 12, 13, 13, 15, 21**

Minimum Value = 1
Q1 = 5
Q2 = Median = 9
Q3 = 13
Maximum Value = 21

## Problem set (2)

1. Find the mean, median, mode, and midrange for the following data set (student's grads in Statistics class):

$$\{70, 82, 61, 100, 32, 49, 0, 88, 93, 76\}$$

2. Find the standard deviation and variance for the following data sets:
   **a)** D1 = {3, 7, 16, 2}
   **b)** D2 = {77, 83, 91, 75, 85, 69}

3. For the following data set: {5, 2, 3, 7, 10, 5, 3, 8, 15, 12, 12, 8}
   **a)** Find the five-Number summary then compute the IQR?
   **b)** Draw a box plot that related to the data points.

4. Find the Maximum and Minimum usual values for a data set:

$$\{8, 14, 12, 6, 10, 16, 4\}$$

5. The mean of statistics final exam is 70 with standard deviation of 6.
   **a)** Find the maximum and minimum usual values?
   **b)** If three students got (81, 85, 61), are these grades considered to be usual or unusual values?

6. At a University, 100 students have been selected and asked "what is your favorite season?". The data has been summarized in the following table. Create a par chart and pie chart for the data set.

| **Season** | **Winter** | **Spring** | **Summer** | **Fall** |
|---|---|---|---|---|
| **Number of students** | 30 | 40 | 10 | 20 |

# الفصل الثالث: مبادى نظرية الاحتمال

**مفاهيم اساسية:**

**التجربة العشوائية (Random Experiment)** : وهي التجربة التي لا يمكن معرفة نتائجها نتيجة لخضوعها لقوانين الاحتمال.

مثال: رمي قطعة من النقود، سحب ورقة من ورقة اللعب، اختيار كرة من صندوق يحتوي على مجموعة كرات ذات ألوان مختلفة.

**فضاء العينة (Sample Space)** : وهي مجموعة كل النتائج الممكنة لتجربة معينة تحدد عن طريق الصدفة ويرمز لها بالرمز {S}.

**مثال1** : عند رمي قطعة من النقود مرة واحدة فأن فضاء العينة هو.

S = {H, T}, where H represents the Head and T represents the Tail.

**مثال2** : عند رمي قطعتين من النقود مرة واحدة فأن فضاء العينة هو.

S = {HH, HT, TH, TT}

**الحادثة (Event)** : وهي مجموعة جزئية من فضاء العينة ويرمز لها بالرمز

{A, B, C, E, E1, E2, ...}

**مثال:** عند رمي قطعة من النقود مرة واحدة فأن فضاء العينة هو: S = {H, T}

E1 = {H}, E2 = {T}, and E3 = {H, T} are events.

**يمكن تصنفين الحوادث الى (Joint and Disjoint Events):**

1. **الحوادث المتنافية (Disjoint Events):** يقال عن الحادثتين E1 وE2 حادثتان متنافيتان إذا استحال حدوثهما معا".

**مثال:** عند رمي قطعة من النقود مرة واحدة فأنه من المستحيل الحصول على H وT في وقت واحد.

2. **الحوادث الغير متنافية (Joint Events):** يقال عن الحادثتين E1 وE2 حادثتان غير متنافيتان إذا أمكن حدوثهما معا".

**مثال:** اختيار شخص ما فأن الحادثتان ان يكون طالب جامعي وموظف هما حادثتان غير متنافيتان.

## يمكن تصنفين الحوادث الى (Dependent and Independent Events):

1 . **الحوادث المستقلة (Independent Events):** هي الحوادث التي وقع أحدهما لا يؤثر في وقوع الاخر.

**مثال:** عند رمي قطعتين من النقود مرة واحدة فأن نتيجة القطعة الأولى لا تؤثر على نتيجة القطعة الثانية.

2 . **الحوادث الغير مستقلة (Dependent Events):** هي الحوادث التي وقوع أحدهما يؤثر على وقوع الحوادث الأخرى.

**مثال:** صندوق يحتوي على كرات بيضاء وسوداء، فعند سحب كرتان على التوالي بدون ارجاع، فأن سحب الكرة الثانية تتأثر بنتيجة السحبة الأولى. بعبارة أخرى نقول ان السحبتان (الحادثتان) غير مستقلة او معتمدة.

---

## بعض العمليات الحسابية:

**المضروب (Factorial)** : مضروب n ويرمز له بالرمز ( **n!** ) ويعرف كما يلي:

$$n! = n*(n-1)*(n-2)*\dots*2*1$$

**ملاحظة:** $0! = 1$ and $n! = n(n-1)!$

**مثال: جد قيمة 3!, 5!, 7! :**

$3! = 3*2*1 = 6$
$5! = 5*4*3*2*1 = 120$
$7! = 7*6*5*4*3*2*1 = 5040$

---

**التباديل (Permutation):** يرمز للتباديل بالرمز ( $P_r^n$ ) ويعرف كالاتي:

$$P_r^n = \frac{n!}{(n-r)!}$$

**مثال:** جد قيمة $P_6^7$ ، $P_1^5$ ، $P_2^3$ :

$$P_2^3 = \frac{3!}{(3-2)!} = \frac{6}{1} = 6$$
$$P_1^5 = \frac{5!}{(5-1)!} = \frac{5.4!}{4!} = 5$$
$$P_6^7 = \frac{7!}{(7-6)!} = \frac{7!}{1!} = 7!$$

**التوافيق (Combination):** يرمز للتوافيق بالرمز ( $\mathbf{C_r^n}$) ويعرف كالاتي:

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

**مثال:** جد قيمة $\mathbf{C_{99}^{100}}$ ، $\mathbf{C_3^6}$ ، $\mathbf{C_2^4}$ :

$$C_2^4 = \frac{4!}{2!\,(4-2)!} = \frac{4.3.2!}{2!\,2!} = 6$$

$$C_3^6 = \frac{6!}{3!\,(6-3)!} = \frac{6!}{3!\,3!} = \frac{6.5.4.3!}{3!\,3!} = 20$$

$$C_{99}^{100} = \frac{100!}{99!\,(100-99)!} = \frac{100.99!}{99!\,.1!} = 100$$

---

**قوانين العد (Counting Rules):**

1. **الطرق الأساسية:** اذا كان عدد الطرق الممكنة لوقوع الحادثة E1 هو(n) و عدد الطرق الممكنة لوقوع الحادثة E2 هو (m) وكانت E1 و E2 حادثتان مستقلتان فأن عدد الطرق الممكنة لوقوع E1 او E1 هو (n*m) من المرات.

**مثال:** إذا كان لشخص أربعة قمصان وستة بنطلون، فكم طريقة يمكن للشخص اختيار زي مختلف؟

الحل: n = 6, m= 4

عدد الطرق = m*n= 6*4=24

---

2. **طريقة التباديل (Permutation Rule):** هو ترتيب r من العناصر مأخوذة من n من العناصر في شكل معين (الترتيب مهم) ويرمز له ب الرمز ( $\mathbf{P_r^n}$) ويعرف كالاتي:

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

where **n** is the total number and **r** is the number that we want to select

**مثال 1 :** إذا كان لدينا مجموعة متكونة من ستة ارقام {1,2,3,4,5,6} فبكم طريقة يمكن تكوين رقم يتكون من ثلاث مراتب؟

الحل: n = 6, r= 3، الترتيب مهم

$$P_r^n = P_3^6 = \frac{6!}{(6-3)!} = \frac{6!}{3!} = \frac{6.5.4.3!}{3!} = 120$$

**مثال 2 :** بكم طريقة يمكن تكوين لجنة تتكون من ثلاث مناصب (الرئيس، النائب وامين السر) اذا علم ان عدد المشاركين بالانتخابات هم عشرة اشخاص؟

**الحل:** n = 10, r= 3، الترتيب مهم

$$P_r^n = P_3^{10} = \frac{10!}{(10-3)!} = \frac{10!}{7!} = \frac{10.9.8.7!}{7!} = 720$$

---

**3. <u>طريقة التوافيق (Combination Rule):</u>** هو عدد طرق اختيار r من العناصر من n من العناصر بدون ترتيب (الترتيب غير مهم)، ويرمز لها بالرمز ( $\mathbf{C_r^n}$) ويعرف كالاتي:

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

where **n** is the total number and **r** is the number that we want to select

**مثال 1:** بكم طريق يمكن تكوين لجنة من أربعة طلاب في صف يحتوي على عشرون طالب؟

**الحل:** n = 20, r= 4 ، الترتيب غير مهم

$$C_r^n = C_4^{20} = \frac{20!}{4!\,(20-4)!} = \frac{20!}{4!\,16!} = \frac{20.19.18.17.16!}{4.3.2.1 * 16!} =$$

**مثال2:** بكم طريقة يمكن تكوين كلمة تتكون من ثلاثة حروف من بين حروف كلمة ذي قار؟

الحل: n = 5, r= 3 ، الترتيب مهم

$$C_r^n = C_3^5 = \frac{5!}{3!\,(5-3)!} = \frac{5!}{3!\,2!} = \frac{5.4.3.2.1}{3.2.1 * 2.1} = 10$$

---

**<u>الاحتمالية</u>** (**Probability**): هي دالة منطقها مجموعة جزئية من فضاء العينة (الحوادث) ومستقرها مجموعة جزئية من الاعداد الحقيقية.

$$\mathbf{P(E)} = \frac{\textbf{Number of ways A can occure}}{\textbf{Number of different simple events}}$$

$$\mathbf{i.e.,} \qquad \mathbf{P(E)} = \frac{\textbf{Part}}{\textbf{Total number}}$$

---

**<u>Remarks:</u>**

1. $0 \leq P(E) \leq 1$
2. If P(E ) = 1, then E is a certain event.
3. If P(E ) = 0, then E is impossible event.

---

**<u>ملاحظة</u>:** يرمز لمكملة (متممة) الحادثة E بالرمز $E^c$. حيث ان:

$$P(E^c) = 1 - P(E) \Rightarrow P(E^c) + P(E) = 1$$

---

**مثال: تم رمي قطعة نقود ثلاث مرات، جد احتمالية وجود صورتان؟**

الحل: فضاء العينة هو:

S = {HHH, HHT, HTH, THH, HTT, THT, TTH, TTT}

Event = E = {HHT, HTH, THH}

$$P(E) = \frac{\text{Part}}{\text{Total}} = \frac{3}{8}$$

**Example:** A box has 3 white balls, 5 black balls, and 2 red balls. If someone select a ball randomly, what is the probability that the ball is (red, not black, not red).

**Solution:**

Total number = 3 + 5+ 2 = 10 balls

E1 = black ball → P(E1) = P(black) = $\frac{5}{10} = \frac{1}{2}$

E2 = red ball → P(E2) = P(red) = $\frac{2}{10} = \frac{1}{5}$

E3 = white ball → P(E3) = P(white) = $\frac{3}{10}$

Hence,

P(black) = $\frac{5}{10} = \frac{1}{2}$

P(not black) =1- $P(\text{ black}) = 1 - \frac{1}{2} = \frac{1}{2}$

P(not white) =1- $P(\text{ white}) = 1 - \frac{3}{10} = \frac{7}{10}$

**Note:** P(black) + P(red) + P(white) = $\frac{\mathbf{5}}{\mathbf{10}} + \frac{\mathbf{2}}{\mathbf{10}} + \frac{\mathbf{3}}{\mathbf{10}} = \mathbf{1}$

---

**Example:** A box has 6 white balls, 4 black balls, and 5 red balls. If someone randomly select a ball, what is the probability that the ball is (white, not white).

**Solution:**

Total number = 6 + 4 + 5 = 15 balls

P(white) $= \frac{9}{15}$

P(not white) =1- $P(\text{ white}) = 1 - \frac{6}{15} = \frac{9}{15}$

**Note: P(white) + $\mathbf{P(\text{ not white}) = 1}$**

---

## قوانين الاحتمالية (Law of Probability):

### 1. قانون الجمع (Adding Rule):

**i.** اذا كانت الحوادث متنافية، أي أذا كانت E1 و E2 حادثتان متنافيتان فان احتمال حدوث أي منهما ( E1 او E2) هو حاصل جمع احتمال كل منهما أي:

$$\mathbf{P(E1 \text{ OR } E2) = P(E1 \cup E2) = P(E1) + P(E2)}$$

**Example:** A box has 6 white balls, 3 black balls, and 5 blue balls. If someone select a ball randomly, what is the probability that the ball is white or black?

**Solution:**

Total number = 6 + 3 + 5 = 14 balls

P(white or black) = P(white) + P(black) = 6/14 +3/14 = 9/14

---

**ii.** اذا كانت الحوادث غير متنافية، اي اذا كانت E1 و E2 حادثتان غير متنافيتان فان احتمال حدوث أي منهما ( E1 او E2) هو حاصل جمع احتمال كل منهما مطروحا" منه احتمال حدوثهما معا" أي:

$$\mathbf{P(E1 \text{ OR } E2) = P(E1 \cup E2) = P(E1) + P(E2) - P(E1 \cap E2)}$$

**Example:** In a class, 25% of the students pass exam1,15% of the students pass exam2, and 10% of the students pass both exam1 and exam2. If someone randomly select a student, what is the probability that the student pass exam1 or exam2?

**Solution:**

P(student pass exam1 or exam2) = P(student pass exam1) + P(student pass exam2)

– p(student pass exam both exam1 and exam2)

→P(student pass exam1 and exam2) = 0.25 + 0.15 - 0.10 = 0.30

### 2. قانون الضرب (Multiplication Rule)

**a)** اذا كانت الحوادث مستقلة، اي اذا كانت E1 و E2 حادثتان مستقلتان فان احتمال حدوثهما معا" هو حاصل ضرب كل منهما اي:

$$\mathbf{P(E1 \cap E2) = P(E1) * P(E2)}$$

**Example:** A box has 6 white balls, 3 black balls, and 5 blue balls. If someone select a ball randomly, what is the probability that the ball is white and black?

**Solution:**

Total number = 6 + 3 + 5 = 14 balls

P(white and black) = P(white) * P(black) = 6/14 *3/14 = 18/196

---

**b)** اذا كانت الحوادث غير مستقلة، اي اذا كانت E1 و E2 حادثتان غير مستقلتان فان احتمال حدوثهما معا" هو:

$$\mathbf{P(E1 \cap E2) = P(E1) * P(E/E1)}$$

---

### الاحتمال الشرطي (Conditional Probability)

اذا كان الحادثة E1 قد وقعت فعلا" والمطلوب إيجاد احتمال وقوع الحادثة E2 بحيث ان E2 يعتمد على E1 فأن الاحتمال الشرطي هو:

$$\mathbf{P(E2/E1) = \frac{P(E1 \cap E2)}{P(E1)}}\text{, where } P(E1) > 0$$

---

**Example:** In a small city, someone wants to select a person from 900 people where these people ( Boy & Girl) and (Job & No job). Find the probability for the following:

1. P(E)
2. P(B)
3. P(G and V)
4. P(B or E)
5. P(B/V)
6. P(G/E)
7. P(E/G)
8. P(V/B)

| | Job or No job | |
|---|---|---|
| **Gender** | **Job: E** | **No job: V** |
| **Boy : B** | 460 | 40 |
| **Girl : G** | 140 | 260 |

**Solution**: _

| Gender | Job: E | No job: V | summation |
|---|---|---|---|
| **Boy : B** | 460 | 40 | 500 |
| **Girl : G** | 140 | 260 | 400 |
| **Summation** | 600 | 300 | **900** |

**1.** P(E) =

**2.** P(B) =

**3.** P(V) =

**4.** P(E) =

**5.** P(G and V) =

**6.** P(B or E) =

**7.** P(B/V) =

**8.** P(G/E) =

**9.** P(E/G) =

**10.** P(V/B) =

# Problem set (3)

1. Give an example for:
   a. Joint and disjoint events.
   b. Dependent and independent events.

2. In how many ways someone can answer five questions from eight questions?

3. Let S = {1,2,3,4,5} be a set of five numbers. In how many ways someone can create a number that consist from three digits such that:
   a. The number is even.
   b. The number is odd.

4. In a classroom, someone wants to select a student from 100 students as described in the following table. Find the following:

| | **Use a drag** | **Does not use a drag** |
|---|---|---|
| **Sick** | 20 | 15 |
| **Non-Sick** | 25 | 40 |

**a)** P(sick)

**b)** P(Does not use a drag)

**c)** P(Non-sick **AND** use a drag)

**d)** P(Sick **OR** Use a drag)

**e)** P(sick **AND** does not use a drag)

**f)** P(Non-sick **/** Use a drag)

**g)** P(Does not use a drag **/**Sick)

**h)** P(sick **/** Use a drag)

**i)** P(Non-sick **OR** Does not use a drag)