



المحاضرة الأولى

الإحصاء الحيوي

السنة الثالثة صيدلة

د. وديع علي

البيانات الإحصائية ومؤشراتها الوصفية

مقدمة:

لقد أصبح علم الإحصاء أداة ضرورية لكل العلوم ووسيلة أساسية لجميع البحوث العلمية، ولهذا فإننا سنقدم لطلابنا في هذه المحاضرات جرعة بسيطة لمبادئ هذا العلم. آملين أن يستوعبونها وأن يستفيدوا منها في حياتهم العملية وفي دراساتهم العلمية.

يعتمد علم الإحصاء على جمع المعلومات الإحصائية عن الظواهر المدروسة ثم معالجتها وتحليلها واستخلاص النتائج منها.

تعريف علم الإحصاء

علم الإحصاء Statistics هو العلم الذي يهتم بالطرق العلمية لجمع وتنظيم وتلخيص البيانات وعرضها وتحليلها بأساليب علمية للحصول منها على المعلومات اللازمة لاتخاذ القرارات المناسبة.

تحليل البيانات أي المعالجة الإحصائية للبيانات قد تتم بطريقة يدوية Manual Processing أو بطريقة الكترونية Electronic Processing باستخدام الكمبيوتر أو بمزيج منهما.

أقسام علم الإحصاء

ينقسم علم الإحصاء إلى قسمين أساسيين هما:

1- الإحصاء الوصفي Descriptive statistics

2- الإحصاء الاستدلالي Statistical inference

الإحصاء الوصفي: هو عبارة عن الطرق العلمية التي تستخدم لتنظيم وعرض ووصف البيانات باستخدام الجداول والرسوم البيانية وبعض المقاييس الإحصائية البسيطة التي تستخدم لتلخيص البيانات (أي أن الإحصاء الوصفي يهتم بجمع وتبويب البيانات الإحصائية).

الإحصاء الاستدلالي: هو عبارة عن الطرق العلمية التي تستخدم للاستدلال عن معالم المجتمع بناءً على المعلومات التي تم الحصول عليها من العينة المأخوذة من هذا المجتمع، وذلك وفق الطرق الإحصائية المعروفة (أي أن الإحصاء الاستدلالي يهتم باستخلاص النتائج من العينة وتعميمها على المجتمع واستخدامها في اتخاذ القرارات المناسبة).

البيانات وأنواعها

تعريف البيانات (Data):

عبارة عن مجموعة من القيم أو المشاهدات أو الملاحظات أو القياسات التي يتم جمعها من مفردات المجتمع أو العينة لخاصية معينة (متغير). وهذه البيانات إما أن تكون كمية (أعداد) أو وصفية.

أنواع البيانات:

تنقسم البيانات إلى نوعين هما:

1- بيانات كمية (Quantitative Data)

2- بيانات وصفية (Qualitative Data)

يمكن تصنيف البيانات الكمية إلى نوعين هما:

• البيانات الكمية المنفصلة: وهي البيانات التي تكون منفصلة ومتباعدة أي البيانات التي تدل على صفة يمكن عدّها (وهذه البيانات تأخذ أعداد صحيحة فقط) مثل أعداد الطلاب، عدد حوادث السيارات، عدد الأسهم الخ...

• البيانات الكمية المتصلة: وهي البيانات التي تكون متصلة ومتقاربة أي البيانات التي لا يوجد لها واقعياً قيمة واحدة حيث تأخذ أي قيمة تقع في مدى معين. أي البيانات التي تدل على صفة يمكن قياسها (وهذه البيانات تأخذ قيماً صحيحة وكسرية) مثل: أوزان الأطفال، درجات الحرارة الخ...

يمكن تصنيف البيانات الوصفية إلى نوعين هما:

• بيانات وصفية اسمية: وهي بيانات تعتمد على التصنيف الوصفي بغض النظر عن أهمية الترتيب مثل زمرة الدم، بيانات النوع (ذكر – أنثى) ، الحالة الاجتماعية (أعزب – متزوج) ، تصنيف موظفي الشركات حسب الجنسية أو حسب التخصص الخ....

• بيانات وصفية ترتيبية: وهي البيانات التي يمكن ترتيبها تصاعدياً أو تنازلياً أي البيانات التي يلعب الترتيب دوراً أساسياً في تحديد معالم الظاهرة محل الدراسة مثل: درجة الألم، مراحل السرطان ، درجة الموافقة على قرار (غير موافق على الإطلاق- غير موافق- محايد – موافق – موافق بشدة)، ترتيب موظفي إحدى الشركات حسب المؤهل العلمي (ثانوي – دبلوم فني – مؤهل عالي – ماجستير – دكتوراه).

ملاحظة: يمكن تحويل البيانات الكمية إلى بيانات نوعية أما العكس، غير صحيح.

مصادر البيانات

يمكن جمع بيانات الدراسة أو البحث من مصدرين هما:

1- المصادر التاريخية: وهي المصادر التي نحصل منها على البيانات بشكل غير مباشر. ويتم الحصول عليها بواسطة أشخاص آخرين أو أجهزة وهيئات رسمية متخصصة وتتمثل في البيانات المنشورة والجهازية للاستخدام مثل الوثائق الإحصائية التي تصدرها الهيئات العامة والخاصة والمراجع الموجودة في المكتبات.

2- المصادر الميدانية: وهي المصادر التي نحصل منها على البيانات بشكل مباشر، حيث يقوم الباحث بجمع البيانات من المفردات محل الدراسة أو البحث وذلك بتصميم استمارة استقصاء (استبيان) تتضمن المتغيرات المطلوب قياسها ويتم الحصول على بيانات الاستمارة بطريقة التجربة والمشاهدة أو بالمقابلة الشخصية أو بالبريد العادي أو الإلكتروني أو بأي وسيلة أخرى.

جدولة البيانات

بعد جمع البيانات من المصادر التاريخية أو المصادر الميدانية فإن هذه البيانات تكون أولية غير منظمة عددياً وتصبح دراستها والحصول منها على أية نتائج ، ولذلك دعت الحاجة إلى تنظيم هذه البيانات وتلخيصها ولكي يمكن الاستفادة من هذه البيانات واستيعابها بطريقة أفضل والحصول منها على بعض النتائج الأولية يتم تنسيق هذه البيانات ووضعها في جداول مناسبة.

وتتم عملية جدولة البيانات على مرحلتين متكاملتين هما (التصنيف) و(التبويب).

(أ) تصنيف البيانات:

أهم طرق تصنيف البيانات:

- التصنيف حسب الصفة المميزة مثل (الجنسية أو الحالة التعليمية أو النوع أو العمل)
- التصنيف الزمني (التاريخي) وذلك حسب الأيام و الأشهر أو السنوات
- التصنيف المكاني (الجغرافي) حسب المناطق أو الجهات.

(ب) تبويب البيانات

يتم تبويب البيانات بحصر أو عد الحالات المتشابهة حسب التصنيف المختار من التصنيفات السابقة ووضعها في جداول مناسبة تلخصها وتوضحها وتنسقها في شكل سهل فهمها واستيعابها كما يمكن مقارنتها والرجوع إليها في أي وقت والاستفادة منها في وصف الظاهرة محل الدراسة.

يمكن التبويب بطريقة يدوية في حالة البيانات البسيطة والتي يكون فيها عدد الاستبيانات صغيراً، كما يمكن إجراء التبويب بطريقة آلية باستخدام الكمبيوتر وذلك باستخدام الحزم الجاهزة للبرامج الإحصائية .

جدولة البيانات الوصفية

مثال 1: فيما يلي البيانات الخاصة بمؤهلات العاملين في أحد البنوك :

جامعي	جامعي	جامعي	دبلوم	دكتوراه	دبلوم	دبلوم
جامعي	جامعي	ثانوي	دبلوم	جامعي	دكتوراه	متوسط
دكتوراه	جامعي	دبلوم	دبلوم	ثانوي	دبلوم	دبلوم
جامعي	دبلوم	ثانوي	دبلوم	ثانوي	جامعي	جامعي
ثانوي	جامعي	دبلوم	دبلوم	جامعي	دبلوم	دبلوم

والمطلوب: إنشاء جدول تكراري بسيط يلخص البيانات السابقة.

الحل: لإنشاء جدول تكراري بسيط يلخص البيانات السابقة نتبع ما يلي:

1- إنشاء جدول كما يلي:

المؤهل العلمي	التكرارات المطلقة n_i	التكرارات النسبية $\frac{n_i}{n}$
دكتوراه	3	0.09
جامعي	12	0.34
دبلوم	14	0.4
ثانوي	5	0.14
متوسط	1	0.03
المجموع	35	1

جدولة البيانات الكمية

تم عملية جدولة البيانات الكمية بتقسيم المجموعة العددية التي لدينا إلى مجموعات جزئية بحيث تحتوي كل مجموعة جزئية عدداً من القيم المتقاربة يطلق على هذه المجموعة الجزئية بالفئة ويفترض أن جميع المفردات التي تقع داخل الفئة متساوية ومساوية لمركز هذه الفئة ثم نحصر العدد داخل كل فئة ويسمى بالتكرار وبذلك تتحول المجموعة العددية الأصلية إلى عدد من الفئات وتكراراتها. وتسمى هذه العملية بالتوزيع التكراري وعند وضعها في جدول تسمى بجدول التوزيع التكراري ((Frequency Distribution أو بالجدول التكراري.

مثال 2: الجدول الآتي يوضح البيانات الخاصة بأجور 50 عاملاً في إحدى الشركات بآلاف الليرات:

22	34	12	36	5	37	28	32	40	25
37	37	37	35	42	30	29	36	40	32
20	33	3	45	24	27	32	35	29	20
15	17	38	30	30	28	10	16	26	25
35	26	26	20	35	26	36	36	43	35

قسم هذه العينة إلى تسع فئات متساوية الطول بحيث تبدأ الفئة الأولى من 3 وعين تكرارات هذه الفئات وتكراراتها النسبية، والتكرارات المئوية وعين التكرار المتجمع الصاعد K_i والتكرار المتجمع الهابط k_i .

الحل: لتكوين جدول التكراري البسيط نتبع ما يلي:

1- حساب المدى المطلق للبيانات:

$$\text{Range} = \text{Maximum} - \text{Minimum}$$



جامعة
المنارة
MANARA UNIVERSITY

$$\text{المدى} = 45 - 3 = 42$$

2- تحديد عدد الفئات:

يتم تحديد عدد الفئات من قبل الشخص القائم بجدولة البيانات ويتحدد عدد الفئات وفقاً لاعتبارات منها: حجم العينة، الهدف من البحث، رأي الباحث ويرى كثير من الباحثين أن عدد الفئات يجب ألا يقل عن 5 فئات ولا يزيد عن 15 فئة حتى لا يكون عدد الفئات صغيراً وبالتالي نفقد كثيراً من التفاصيل وتضيع معالم التوزيع، كما يجب ألا يكون عدد الفئات كبيراً فتضيع الحكمة من التجميع في فئات وبالتالي نقوم باختيار عدد الفئات وليكن 9 فئات

3- حساب طول الفئة:

طول الفئة = (المدى : عدد الفئات)

$$L = \frac{42}{9} = 4.66 \approx 5$$

4- جدول التوزيع التكراري البسيط:

مجال الفئة I	التكرار n_i	التكرار النسبي $\frac{n_i}{n}$	التكرار المئوي $\frac{n_i}{n} \times 100$	التكرار الصاعد K_i	التكرار الهابط K'_i
[3,8[2	0.04	4%	2	50
[8,13[2	0.04	4%	4	48
[13,18[3	0.06	6%	7	46
[18,23[4	0.08	8%	11	43
[23,28[8	0.16	16%	19	39
[28,33[10	0.2	20%	29	31
[33,38[15	0.3	30%	44	21
[38,43[4	0.08	8%	48	6
[43,48[2	0.04	4%	50	2
المجموع	50	1	100%	-	

العرض البياني للبيانات:

في المرحلة السابقة تكلمنا عن عرض البيانات جدولياً أي تنظيمها في جداول تسهل من فهمها واستيعابها، وقد لاحظنا أن العرض الجدولي للبيانات يعطي صورة واضحة عن البيانات الأولية وتوزيعاتها التكرارية ومع ذلك فإن العرض البياني للبيانات يعطي فكرة أسرع وأوضح عن أشكال التوزيعات التكرارية للبيانات وبالتالي نتناول عرض البيانات بيانياً حسب أنواع البيانات.

مثال 3:

حصلنا من سجلات مديرية الصحة في محافظة اللاذقية على الدخل الأسبوعي بالآلاف لخمسين صيدلاني أخذوا بشكل عشوائي. وكانت هذه الدخول موزعة كما يلي:

111	81	129	87	98	91	86	102	96	84
106	81	127	74	91	82	94	96	94	96
94	83	96	71	96	115	83	109	87	96
107	89	118	80	121	98	83	103	86	86
118	126	110	107	104	120	123	102	114	130

والمطلوب:

قسم هذه العينة إلى سبع فئات متساوية الطول بحيث تبدأ الفئة الأولى من 65 وعين مراكز وتكرارات هذه

الفئات وتكراراتها النسبية، وعين التكرار المتجمع الصاعد K_i والتكرار المتجمع الهابط \hat{k}_i .

الحل:

طول الفئة: بما أن المدى لهذه العينة هو $130 - 71 = 59$ وبما أن المطلوب تقسيمها إلى سبع فئات متساوية الطول:

$$L = \frac{130-65}{7} \cong 9.29$$

ولذلك نختار طول الفئة يساوي 10 ويكون لدينا الجدول التكراري الآتي:

رقم الفئة	مجال الفئة I	مركز الفئة x_i	تكرار الفئة n_i	التكرار النسبي للفئة	k_i	\hat{k}_i
1	[65-75[70	2	0.04	2	50
2	[75-85[80	8	0.16	10	48
3	[85-95[90	11	0.22	21	40
4	[95-105[100	12	0.24	33	29
5	[105-115[110	7	0.14	40	17
6	[115-125[120	6	0.12	46	10
7	[125-135[130	4	0.08	50	4

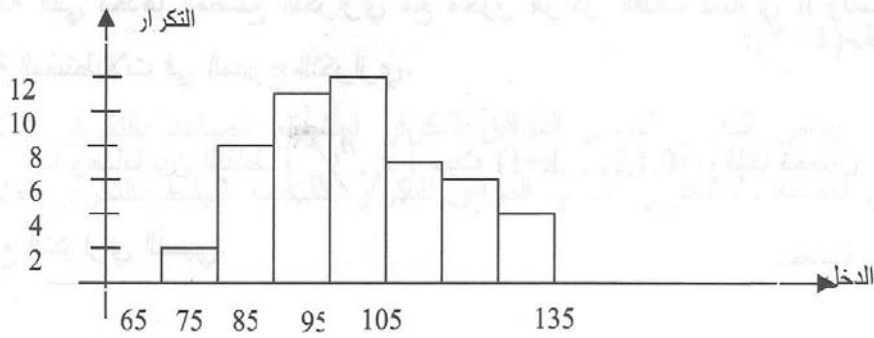
4-1 التمثيل البياني لعينة:

يمكن تمثيل جداول التوزيعات التكرارية بيانياً بأشكال مختلفة أهمها:

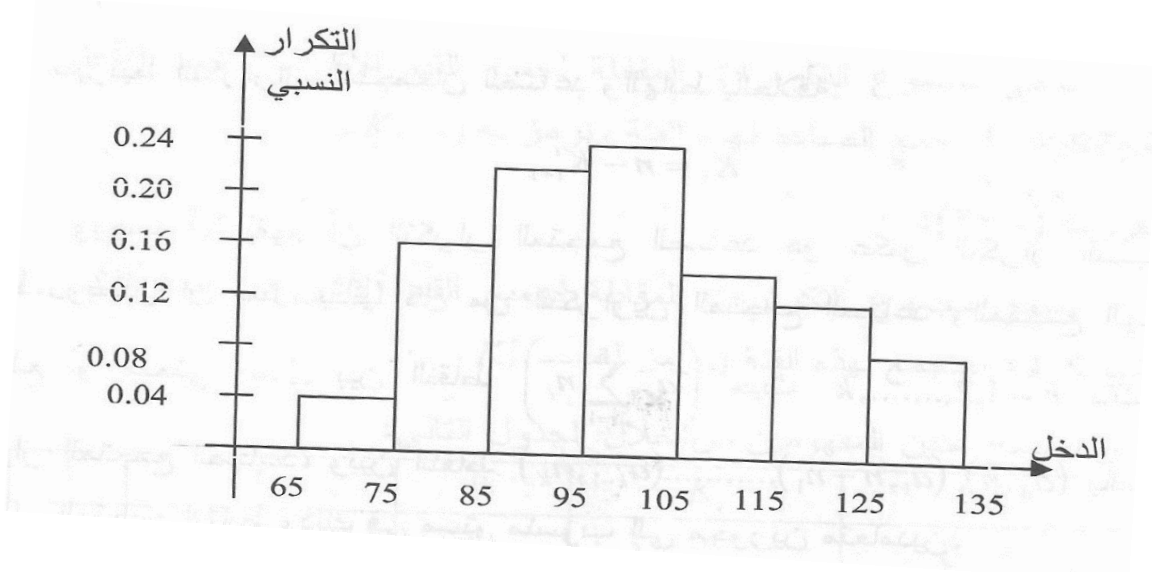
أ- المدرج التكراري: (Frequency histogram)

يرسم في مستوٍ منسوب إلى محورين متعامدين (حيث المحور الأفقي هو محور مجالات الفئات والمحور الشاقولي محور التكرارات) مستطيلات متلاصقة قواعدها منطوية على مجالات الفئات I_i وارتفاعاتها تساوي تكرارات الفئات n_i (وإذا كانت من أجل $i = 1, 2, \dots, k$ ارتفاع المستطيلات تساوي التكرارات النسبية $\frac{n_i}{n}$ من أجل $i = 1, 2, \dots, k$ فإننا نحصل على المدرج التكراري النسبي للعينة).

بالعودة إلى المثال (3) فإن المدرج التكراري للعينة يأخذ الشكل:



والمدرج التكراري النسبي للعينة:

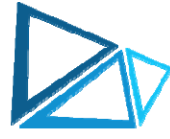


ملاحظة:

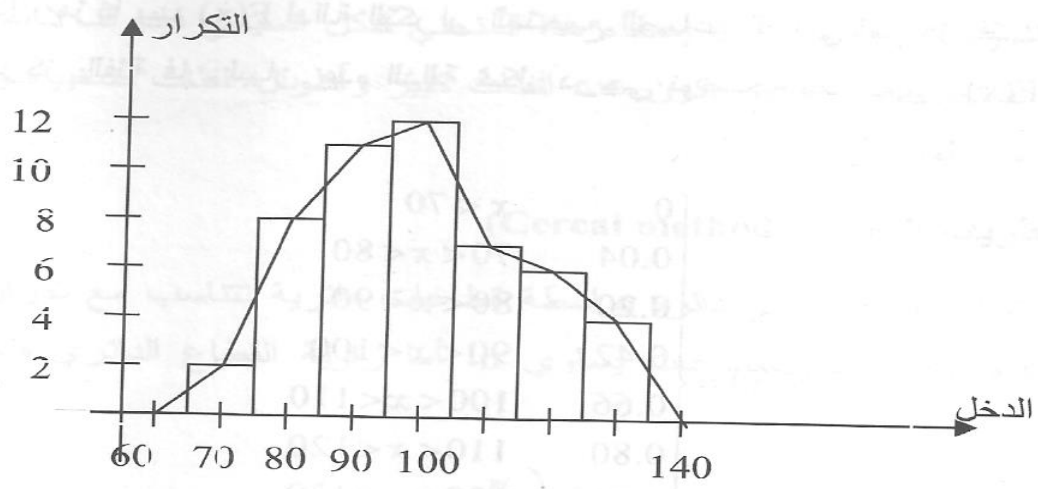
إذا اعتبرنا أن طول مجال الفئة يساوي واحدة الأطوال فإن مجموع مساحات المستطيلات في المدرج التكراري يساوي n وفي المدرج التكراري النسبي يساوي الواحد.

ب- المضلع التكراري لعينة (Frequency polygon)

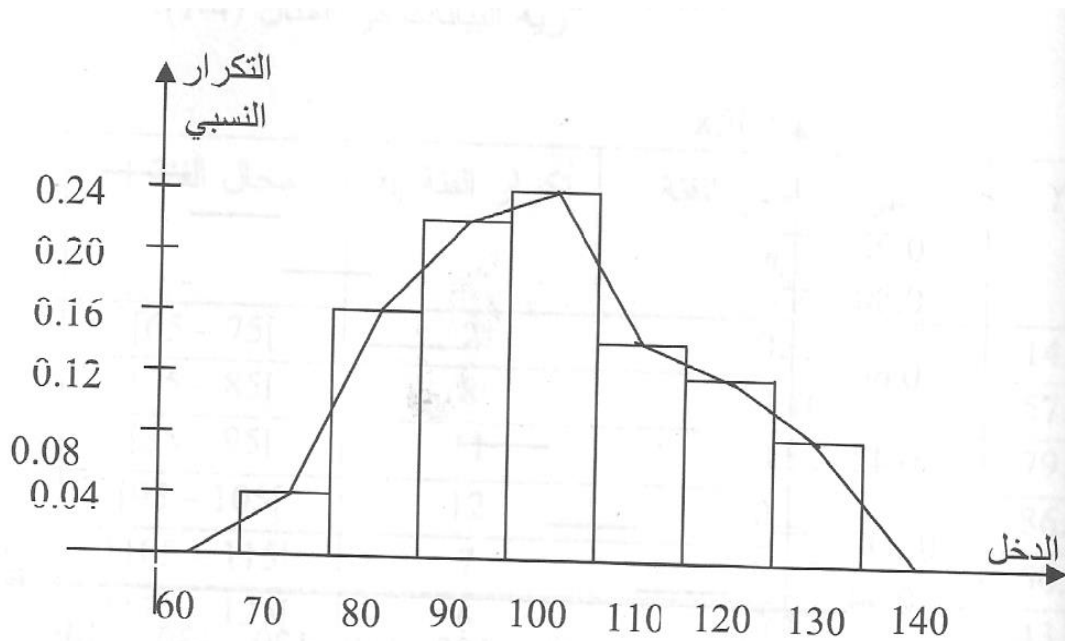
يرسم في مستوٍ منسوب إلى محورين متعامدين (حيث المحور الأفقي هو محور مراكز الفئات والمحور الشاقولي محور التكرارات) النقاط (x_i, n_i) حيث x_i مركز الفئة i و n_i تكرار الفئة i حيث يكون الخط المنكسر الذي يصل بين كل نقطتين متتاليتين من هذه النقاط $i = 1, 2, \dots, k$ و عدد الفئات ويضاف لمجموعة هذه النقاط النقطتين $(x_0, 0)$ و $(x_{k+1}, 0)$ فيكون الخط المنكسر الذي يصل بين كل نقطتين متتاليتين من النقاط $(x_0, 0)$ و $(x_{k+1}, 0)$ هو المضلع التكراري للعينة وإن المساحة التي يحددها المضلع التكراري مع محور مراكز الفئات تساوي n وتساوي مجموع المساحات المستطيلات في المدرج التكراري. وإذا وصلنا بين النقاط $(x_i, \frac{n_i}{n})$ حيث $i = 1, 2, \dots, k + 1$ فإننا نحصل على المضلع التكراري النسبي. يأخذ المضلع التكراري للعينة في المثال (3) الشكل:



جامعة
المنارة
MANARA UNIVERSITY



والمضلع التكراري النسبي:



ملاحظة:

يسمى التكرار النسبي الموافق للتكرار المتجمع الصاعد بالتكرار المتجمع النسبي الصاعد. والتكرار النسبي الموافق للتكرار المتجمع الهابط بالتكرار المتجمع النسبي الهابط.

ج- المنحني التكراري لعينة: (Frequency curve)

هو المنحني المار من النقاط (x_i, n_i) من أجل $i = 0, 1, 2, \dots, k + 1$ في مستوي منسوب إلى محورين متعامدين، حيث المحور الأفقي هو محور مراكز الفئات والمحور الشاقولي هو محور تكرارات مراكز الفئات

ويسمى المنحني المار بالنقاط $\left(x_i, \frac{n_i}{n}\right)$ بالمنحني التكراري النسبي حيث $i = 0, 1, 2, \dots, k + 1$.

تجدر الإشارة إلى أنه من أجل إعطاء هذا المنحني شكل انسيابي فإنه قد لا يمر ببعض النقاط ويستخدم عادة عندما يكون عدد الفئات كبير وأطوال الفئات صغيرة وهذا يبين عدم دقة هذا المنحني.

د- طريقة الدائرة (Circle method)

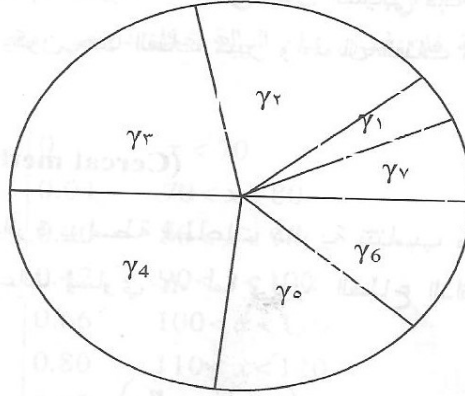
تمثل البيانات على دائرة بواسطة قطاعات دائرية تتناسب مع تكرار الفئات ومجموع مساحات هذه القطاعات

يساوي n أما زاوية القطاع الدائري فتحسب من العلاقة: $\gamma_i = 360 \frac{n_i}{n}$; $i = 0, 1, 2, \dots, k$

لنمثل بيانياً بواسطة قطاعات دائرية البيانات في المثال (3):

رقم الفئة	مجال الفئة I	تكرار الفئة n_i	التكرار النسبي للفئة	γ_i
1	[65-75[2	0.04	14.4
2	[75-85[8	0.16	57.6
3	[85-95[11	0.22	79.2
4	[95-105[12	0.24	86.4
5	[105-115[7	0.14	50.4
6	[115-125[6	0.12	43.2
7	[125-135[4	0.08	28.8

عندئذ:



مثال (4):

تمثل القراءات التالية الزيادات في الوزن بالمقارنة مع الطول لعينة من عشرين شخص:

5.3	9.2	1.2	5.8	0.8	6.1	2.6	7.4	6.2	3.3
6.5	3.4	7.1	7.3	3.7	7.8	4.0	2.8	7.6	4.5

- 1- رتب هذه العينة تصاعدياً.
- 2- رتب هذه العينة في جدول من خمس فئات بحيث تبدأ الفئة الأولى من الصفر، مبيناً التكرار النسبي – والتكرار المتجمع الصاعد – والتكرار المتجمع الهابط.
- 3- ارسم المدرج والمضلع التكراريين لهذا البيان الإحصائي.

الحل:

1-

0.8	1.2	2.6	2.8	3.3	3.4	3.7	4.0	4.5	5.3
5.8	6.1	6.2	6.5	7.1	7.3	7.4	7.6	7.8	9.2

- 2- بما أن الفئة الأولى تبدأ من الصفر وأكبر عدد في العينة هو 9.2 وبما أن عدد الفئات خمس فإن طول

$$\frac{9.2}{5} \cong 1.84 \text{ الفئة يساوي}$$

ولذلك نختار طول الفئة 2

i	I_i	x_i	n_i	n_i / n	K_i	K_i'
1	[0-2[1	2	2/20	2	20
2	[2-4[3	5	5/20	7	18
3	[4-6[5	4	4/20	11	13
4	[6-8[7	8	8/20	19	9
5	[8-10[9	1	1/20	20	1

-3

