المحاضرة الخامسة

جداول التوزيع التكراري

يقصد بجدول التوزيع التكراري تنظيم القيم الإحصائية ضمن فئات متسلسلة، بحيث يظهر أمام كل فئة عدد مرات تكرارها، أو بمعنى آخر على سبيل المثال تصنيف البيانات الجغرافية حسب تكرار الظاهرة المدروسة.

لا بد لفهم وإعداد جداول التوزيع التكراري من فهم المراحل التالية:

1- ترتيب القيم تصاعدياً أو تنازلياً:

على سبيل المثال مردود هكتار واحد من القمح في سورية خلال الفترة بين الأعوام (1950 – 1969) يتم ترتيبها تصاعدياً كما في الجدول التالي:

جدول 1

رقم التسلسل	المردود	رقم التسلسل	المردود
1	115	11	685
2	300	12	716
3	357	13	745
4	387	14	763
5	445	15	822
6	490	16	822
7	574	17	859
8	654	18	874
9	662	19	915
10	673	20	969

مثال:

لدينا بيانات حول أعمار مجموعة من الأشخاص عددهم 30 شخص في محافظة دمشق ونريد تبويب هذه البيانات (المتغيرات) في جدول تكراري:

19	20	23	32	30	15	19	27
20	26	27	21	22	19	18	22
20	16	20	18	15	11	09	26
		17	21	21	25	21	20

- في البداية نرتب المعلومات الواردة في الجدول ترتيباً تصاعدياً لتصبح كما يلي:

09	11	12	15	15	16	17	18
18	19	19	20	20	20	20	20
20	21	21	22	22	23	24	25
26	26	27	27	30	32		

- الخطوة الثانية نخفض حجم البيانات وتلخيصها في جدول جديد نسميه "جدول التوزيع التكراري": يتم في هذا الجدول تنظيم المفردات في فئات متسلسلة، بحيث يظهر أمام كل فئة عدد مرات تكرارها.

جدول 2

فئات العمر	عدد الأشخاص
09 – 13	3
14 – 18	6
19 – 23	13
24 – 28	6
29 – 33	2
المجموع	30

بعض المفاهيم الأولية للتوزيع التكراري:

- الفئات: هي مجموعات متقاربة من البيانات الإحصائية للمتغيرات يتم تقسيمها بحيث يكون لكل فئة حدها الأعلى وحدها الأدنى مع تحديد التكرارات المقابلة لكل فئة.
- المدى العام: هو القيمة العددية بين أعلى قيمة وأدنى قيمة في مجمل البيانات الإحصائية. أي بمعنى آخر هو الفرق بين القيمة الأخيرة والقيمة الأولى في البيانات بعد ترتيبها ترتيباً تصاعدياً.
- عدد الفئات: هناك علاقات يمكن أن تعطينا الحد الأقصى لعدد الفئات المناسبة التي يتكون منها الجدول التكراري، وهي ليست علاقات قطعية إنما تعطي مؤشرات للباحث ليختار الأنسب من مجمل مفردات بياناته. ومنها العلاقة الآتية:

$$\text{عدد الفئات} = 5 + \log n$$

حيث n : عدد المفردات. (في المثال السابق n هي عدد الأشخاص)

2- تحديد طول الفئات (فترات الفئات):

إن طول الفئة: هو القيمة العددية الواقعة بين حدي الفئة الأدنى والأعلى. فالرقم (9) في المثال السابق يدعى الحد الأدنى للفئة، والرقم (13) يدعى الحد الأعلى للفئة، وبالتالي طول الفئة هنا = (4).

إن من أهم المسائل الواجب حلها عند إعداد جدول التوزيع التكراري هو تحديد طول الفئات حسابياً بشكل مسبق. هذا ما يمكننا أن نحصل عليه من خلال الاستعانة بالقانون الآتي المقترح من قبل العالم (سترجس):

$$D = \frac{\text{المدى}}{1 + (3.322 \times \log n)}$$

حيث:
D : طول الفئة.

n: عدد المفردات (في مثالنا السابق n= 30).

بتطبيق هذا القانون على مثالنا السابق نجد:

$$D = \frac{32 - 9}{1 + (3.322 \times \log 30)}$$

$$= \frac{23}{1 + (3.322 \times 1.477)}$$

$$= 3.89 \approx 4$$

أي أن طول الفئة هو (4) كما ورد في الجدول السابق.

3- كتابة حدود الفئات:

عند اختيار حدي الفئة يجب أن يراعى عدم التداخل بين الوحدات فمثلا لا يجوز أن تكون الحدود في مثالنا السابق :

$13 - 9$	$25 - 21$
$17 - 13$	$29 - 25$
$21 - 17$	$33 - 29$

لأن في ذلك تداخل بين الفئات، فمثلاً الوحدة الإحصائية ذات القيمة (13) هل تدخل في الفئة الأولى أم في الفئة الثانية ؟ لذلك يفضل كتابة الحدود كما وردت في الجدول 2.

4- حساب التكرارات:

بعد أن يتم تحديد طول الفئات وحدودها، يجري توزيع المفردات في فئات، ويتم ذلك بقراءة الأرقام رقماً رقماً، وتسجيل كل متغير في الفئة التي يعود إليها، حيث توضع إشارة على شكل خط مائل (/) لكل رقم في الفئة التي يقع ضمنها، وكلما أصبح عدد هذه الإشارات خمسة تكتب على شكل حزمة (###) حتى يسهل عدها، وتسمى هذه العملية بعملية التفريغ، وبذلك يجري ترجمة تلك الإشارات على شكل أرقام مقابلة للفئات، في حقل التكرارات كما يلي:

جدول 3

فئات العمر بالسنوات	العلامات	عدد الأشخاص
09 وأقل من 13	///	3
13 وأقل من 17	///	3
17 وأقل من 21	/ ### ###	11
21 وأقل من 25	/ ###	6
25 وأقل من 29	###	5
29 وأقل من 33	//	2
المجموع		30

- **أشكال الجداول التكرارية:**

- **الجداول التكرارية المئوية:**

عندما يكون عدد مفردات الترتيب كبيراً جداً فإنه يمكن إعطاء القارئ صورة أكثر وضوحاً وذلك بتحويل الأرقام التكرارية في كل فئة من الفئات إلى نسبة مئوية من مجموع التكرار العام.

مثال:

عدد سكان مدينة دمشق عام 1970 حسب فئات الأعمار، كما في الجدول الآتي:

جدول 4

فئات العمر	عدد السكان	النسبة المئوية من المجموع العام
أقل من سنة	27714	3.3
4 – 1	112414	13.4
9 – 5	134357	16.1
14 – 10	108659	13.0
19 – 15	81278	9.7
24 – 20	72299	8.6
29 – 25	58512	7.0
34 – 30	51281	6.1
39 – 35	45805	5.5
44 – 40	35892	4.3
49 – 45	37220	3.3
54 – 50	19700	2.4
59 – 55	16662	2.0
64 – 60	15241	1.8
69 – 65	10078	1.2
74 – 70	9304	1.1
79 – 75	4261	0.5
84 – 80	3303	0.4
85 فأكثر	2639	0.3
المجموع	836619	%100.0

*المصدر: نتائج التعداد العام للسكان لعام 1970، رئاسة مجلس الوزراء، المكتب المركزي للإحصاء، ص6.

حسبت النسبة المئوية للفئة الأولى مثلاً كما يلي:

$$\frac{27714}{836619} \times 100 = 3.3\ \%$$

وهكذا لباقي الفئات.

نلاحظ في هذا الجدول الحد الأعلى للفئة الأخيرة غير محدد مما يجعل من الصعوبة تحديد نقطة الوسط لهذه الفئة.

- **الجداول التكرارية التجميعية:**

يسمى الجدول التكراري السابق (4) بجدول تكراري من النوع التجميعي عندما تكون قيم النسبة المئوية من أصل المجموع النهائي تراكمية أي لكل فئة النسبة المئوية الخاصة بها مضافة للنسب المئوية للفئات التي تسبقها، كما في يلي:

جدول 5

فئات العمر	عدد السكان	النسبة المئوية من المجموع العام
أقل من سنة	27714	3.3
4 – 1	112414	16.7
9 – 5	134357	32.8
14 – 10	108659	45.8
19 – 15	81278	55.5
24 – 20	72299	64.1
29 – 25	58512	71.1
34 – 30	51281	77.2
39 – 35	45805	82.7
44 – 40	35892	87.0
49 – 45	37220	90.3
54 – 50	19700	92.2
59 – 55	16662	94.7
64 – 60	15241	96.5
69 – 65	10078	97.5
74 – 70	9304	98.8
79 – 75	4261	99.3
84 – 80	3303	99.7
85 فأكثر	2639	100.0

محاسن الجداول التكرارية:

1. تكون البيانات ملخصة بشكل يسهل عرضها موجزة.
2. يمكننا ملاحظة التشتت الحاصل بنظرة سريعة.
3. يمكن ملاحظة ميل تمركز التكرار حول قيمة معينة.

مساوئ الجداول التكرارية:

1. يمكن أن تفقد المفردات هويتها نظراً لدمج عدة أرقام معاً في فئة واحدة.
2. صعوبة حساب بعض المقاييس الإحصائية كالوسط الحسابي عندما تكون الفئة الأخيرة مفتوحة.

.....