المحاضرة الرابعة

تصنيف وتبويب البيانات

ليس من السهل على الباحث أن يستنتج شيئا من البيانات دون تبويبها [البيانات غير المبوّبة] خاصة عندما تكون الكمية كبيرة، وعملية التعامل مع البيانات من أهم مراحل البحث الإحصائي لذلك يجب وضعها في جدول تكراري.

والتبويب [أي وضع البيانات في جدول] يصلح للبيانات الكمية والوصفية معا.

ويمكن التميز بين شكلين رئيسيين من الجداول الإحصائية، وهما الجدول البسيط الذي يشمل بعد واحد للمتغير، وجدول مزدوج أو متقطع الذي يبوب متغير الذي يبوب متغير واحد أو متغيرين أو أكثر وذلك حسب الظاهرة المدروس.

ويمكن تصميم الجداول حسب المتغير كمي أو نوعي.

1-الجداول التكرارية للبيانات لنوعية:

رصد الفئات أو المجموعات النوعية الممكنة في عمود مع رصد عدد المفردات ويسمى: التكرار المطلق، وهو عدد المرات التي ظهرت فيه صفة معيّنة.

مع مراعاة مجموع الأعداد الذي يجب أن يساوي عدد مفردات العينة.

$$N = \sum_{i=1} n_i = n_1 + n_2 + \cdots + n_k$$

المتغير	التكرار المطلق
الصفة 1 x_2	n_1
الصفة 2 x_2	n_2
. . . الصفةk x_2	. . . n_i
المجموعΣ	N

2-التوزيع التكراري النسبي

التكرار النسبي F_1 لأي فئة أو صفة في الجدول التكراري هو حاصل قسمة تكرار تلك الفئة على مجموع التكرارات الكلي أي:

(مع مراعاة مجموع التكرارات النسبية =1 $F_i = \frac{n_i}{N}$)

$$\sum_i F_i = \sum_i \left(\frac{n_i}{N}\right) = \frac{n_1}{N} + \frac{n_2}{N} + \cdots + \frac{n_k}{N}$$

فالتكرار النسبي هو المقدار الكسري الذي تأخذه كل فئة في الجدول التكراري من المجموع الكلي للتكرارات.

- النسبة المئوية لفئة في الجدول التكراري

نحصل عليها بضرب تكرار تلك النسبة في 100.

3-التوزيع التكراري المتجمع الصاعد:

من أجل فهم هذا العنصر ننطلق من جدول جاهز في فئات وتكرارات، ويكون المطلوب منا إكماله (هي بيانات نتحصل عليها من الدرجات...).

في بعض الأحيان يكون الهدف من الدراسة معرفة عدد أو نسبة المفردات، والتي يتم تصنيفها في أكثر من فئة واحدة.

نأخذ **مثالا:** لدينا البينات التالية:

الفئات	التكرار المطلق n_i	التكرار النسبي	النسبة المئوية	التكرار المتجمع الصاعد $N_i \uparrow$	المتجمع الصاعد النسبي $F_i \uparrow$
ممتاز	10	$^{10}/_{50} = 0.2$	$0.2 \times 100 = 20\%$	10	$^{10}/_{50} = 0.2$
فوق المتوسط	6	$^{6}/_{50} = 0.12$	$0.12 \times 100 = 12\%$	16=10+6	$^{16}/_{50} = 0.32$
متوسط	8	$^{8}/_{50} = 0.16$	$0.16 \times 100 = 16\%$	24=16+8	$^{24}/_{50} = 0.48$
تحت المتوسط	14	$^{14}/_{50} = 0.28$	$0.28 \times 100 = 28\%$	30=24+14	$^{30}/_{50} = 0.76$
ضعيف	12	$^{12}/_{50} = 0.24$	$0.24 \times 100 = 24\%$	50=38+12	$^{50}/_{50} = 1$
المجموع	50	1	100%		

4-التوزيع التكراري المتجمع النازل:

في الحالات التي يكون فيها اهتمامنا منصبا على عدد القيم التي تكون أكبر أو تساوي قيمة معينة.

مثال: إليك البيانات التالية

نقطة الإحصاء	20	13	19	09	10
التكرار	4	3	7	5	1

المطلوب: تكوين جدول التوزيع التكراري المتجمع النازل

قبل البداية نراجع التكرار المتجمع الصاعد في هذه البيانات

الفئات	التكرار	التكرار المتجمع الصاعد
20	4	4
13	+ 3	7
19	+ 7	14
09	+ 5	19
10	+ 1	20
	20	

أما بالنسبة لتكرار المجمع النازل، فيكون كالآتي:

الفئات	**التكرار**	**التكرار المتجمع الصاعد**	النازل هو عكس الصاعد أي نضيف للقيمة الأولى القيم التي بعدها وهكذا
20	4	20= 1+5+7+3+4	
13	+ 3	16	
19	+ 7	13	
09	5	6	
10	1	1	
	20		

5-التوزيع التكراري للفئات:

عندما يكون حجم البيانات كبيرة فيكون من الصعب عرضها بتوزيع تكراري وفي هذه الحالة يلجأ إلى التوزيع التكراري للفئات.

سنعرض مثالا مباشرة لفهم الطريقة، فتقسيم البيانات يكون إلى فئات:

مثال:

75	61	82	57	63	93
88	77	80	69	82	47
81	67	68	76	84	54
94	52	80	94	94	60
66	85	45	76	53	56
71	84	76	59	75	53
		74	57	50	78

1-يجب تحدي أقل قيمة وأكبر قيمة.

أكبرها: 94

أقلها: 45

2-حساب المدى:

$$R = n_{mx} - n_{mn} = 94 - 45 = 49$$
$$R = 49$$

3-حساب عدد الفئات:

$$\log(40) = 1.60$$
$$K = 1 + 3.322\log(n) = 1 + 3.322\log(40)$$

قيمة ثابتة قيمة ثابتة

$$K = 1 + 3.322\log(40) = 1 + 3.322 \times 1.60$$
$$K = 6.34 \cong 6$$

4-حساب طول الفئة:

$$L = \frac{R}{K} = \frac{49}{6} = 8.16$$

$$L \cong 8$$

5-الجدول التكراري:

ملاحظات:

- [المجال المغلق تحتسب القيمة التي تليه.
-] المجال المفتوح لا تحتسب القيمة التي تليه.
- بالنسبة للتكررت نتحصل عليها في الفئات بإحتساب كل الأعداد المحصورة في المجال المذكور.
- المجال المفتوح والمغلق يكون حسب تواجد العدد في البيانات.

الفئات	التكرارات
[45-52]	4
[53-60]	6
[61-68]	6
[69-76]	7
[77-84]	9
[85-92]	2
[93-100]	4

يمكن وضع [93-99[[93-94] لأنها آخر فئة

بعد تكوين الجدول التكراري للفئات نقوم بتكوين المجتمع الصاعد↗ والنازل ↙

1-جدول التكرار المجتمع الصاعد:

الفئات	التكرارات	الحد الأعلى	المجتمع الصاعد $N \nearrow$
[52-45]	4	أقل من 53	4
[60-53]	6	أقل من 61	10
[68-61]	6	أقل من 69	16
[76-69]	7	أقل من 77	23
[84-77]	9	أقل من 85	32
[92-85]	2	أقل من 93	34
]100-93]	4	أقل من 100	38
	مج 38		

ملاحظة: لوكان المجال مفتوح نأخذ القيمة الكبرى، باعتبارها غير محسوبة في المجال.

2-جدول التكرار المجتمع النازل

الفئات	التكرارات	الحد الأعلى	المجتمع الصاعد $N \nearrow$
[52-45]	4	45 فأكثر	38
[60-53]	6	53 فأكثر	34
[68-61]	6	61 فأكثر	28
[76-69]	7	69 فأكثر	22
[84-77]	9	77 فأكثر	15
[92-85]	2	85 فأكثر	6
]100-93]	4	93 فأكثر	4
	مج 38		

تعليمات:

من أجل تكوين توزيع تكراري للفئات يجب:

1. حساب المدى المطلق للبيانات "R"

المدى= أكبر مشاهدة – أصغر مشاهدة

2. تحديد عدد الفئات: "K"

عدد الفئات لا يقل عن 5 ولا يزيد عن 15

$$K = 1 + 3.322\log(M)$$

حجم العينة قيمة ثابتة قيمة ثابتة

3. **نحددّ طول الفئة:** أي مجموع عدد البيانات في كل فئة.

$$L = \frac{R}{K} = \frac{\text{المدى المطلق}}{\text{عدد الفئات}}$$

ملاحظة: يمكن اعتماد عدد الفئات كحدّ أقصى في حالات خاصة "20"

4. اتجاه حدود الفئات والحدود الفعلية ومراكز الفئات

رقم الفئة	حدود الفئات الحد الدنى: الحد الأعلى السابق+1 الحد الأعلى: الحد الأدنى +طول الفئة −1	الحدود الفعلية للفئات الحد الأدنى الفعلي: الحد الأدنى −0.5 الحد الأعلى الفعلي: الحد الأعلى −0.5	مراكز الفئات
1	الحد الأدنى: أصغر مشاهدة الحد الأعلى: الحد الأدنى +طول الفئة −1	الحد الأدنى الفعلي: أصغر مشاهدة 0.5 الحد الأعلى الفعلي: الحد العلى +0.5	$\frac{\text{الحد الأدنى + الحد الأعلى}}{2}$
2	. . . نفس العملية مع اعتماد الحد الأعلى +1، في الفئة السابقة كحد أدنى في هذه الفئة. الحد الأعلى: نفس الخطوات السابقة . . . إلخ	. . . الحد الأدنى الفعلي هو الحد الأعلى الفعلي للفئة السابقة الحد الأعلى: الحد الأدنى الفعلي+ طول الفئة . . . إلخ	

تطبيق المحاضرة 4

اليك البيانات التالية:

الفئات	0−4	5−9	10−14	15−19	20−24	25−29
التكرار	4	5	15	25	6	5

المطلوب:

1-جدول التكرارات النسبية؟

2-جدول التوزيع التكراري المئوي؟

حل تطبيق الحاضرة 4

1-جدول التكرار النسبية:

جدول التوزيع التكرار النسبي مكوّن من عمودي العمود الأوّل هو عمود الفئات، والعمود الثاني هو عمود التكرارات النسبية بحيث يتم استبدال التكرارات العادية بالتكرارات النسبية.

الفئات	التكرارات	التكرار النسبي $= \frac{\text{تكرار الفئة}}{\text{مج التكرارات}}$
4-0	4	$\frac{4}{100} = 0.04$
9-5	5	$\frac{5}{100} = 0.50$
14-10	15	$\frac{15}{100} = 0.15$
19-15	25	$\frac{25}{100} = 0.25$
24-20	6	$\frac{6}{100} = 0.6$
29-25	5	$\frac{5}{100} = 0.05$

2-جدول التوزيع التكراري المئوي

الفئات	التكرارات	التكرار المئوي $=$ التكرار النسبي $\times$ **100**
4-0	4	$0.04 \times 100 = 4\%$
9-5	5	$0.05 \times 100 = 5\%$
14-10	15	$0.15 \times 100 = 15\%$
19-15	25	$0.25 \times 100 = 25\%$
24-20	6	$0.06 \times 100 = 6\%$
29-25	5	$0.05 \times 100 = 5\%$

تقويم:

التمرين الأول:

1- إليك البيانات التالية تظهر عدد السكان لدولة ما حسب فئات الأعمال من 10 إلى 70 سنة

الفئات	[19-10]	[29-20]	[39-30]	[49-40]	[59-50]	المجموع
التكرار	4	9	15	8	4	40

المطلوب:

1. كوّن جدول التوزيع التكراري المتجمع الصاعد.
2. كوّن التكرار المتجمع النازل لنفس البيانات.

التمرين الثاني:

فيمايلي درجات 70 طالبا

56	75	70	56	66	60	55	65	70	65
66	71	62	60	67	71	61	67	61	70
75	69	71	68	57	59	72	68	57	72
65	63	73	72	66	63	58	73	67	62
58	74	60	58	81	80	74	76	74	73
72	94	78	76	91	85	77	83	77	82
62	78	88	79	64	87	55	79	57	64

المطلوب:

1. كوّن التوزيع التكراري لدرجات الطلاب؟
2. كوّن التوزيع التكراري النسبي.
3. ماهي نسبة الطلاب الحاصلين على درجات بين 70 و80.
4. ماهي نسبة الطلاب الحاصلين على درجة أقل من 70.

5. ماهي نسبة الطلاب الحاصلين على درجة 80 أو أكثر

تصحيح التقويم

التمرين الأول:

1. جدول التوزيع التكراري المتجمع الصاعد:

التكرار المجتمع الصاعد:

الفئات	التكرارات F_i	الحد الأعلى	$N\uparrow$
[10−19	4	أقل من 19	4
[20−29	9	أقل من 29	13
30−39	15	أقل من 39	28
40−49	8	أقل من 49	36
50−59	4	أقل من 59	40
	40		

من الجدول أعلاه يمكن معرفة التكرارات التي تقل عن أي حد من حدود الفئات المحددة ويلاحظ أن التجميع يجري بصفة تصاعدية أي من الأدنى إلى الأعلى. لهذا يسمى هذا التجميع بالتوزيع التكراري المتجمع الصاعد ويرمز للتكرارات المتجمعة الصاعدة بسهم نحو الأعلى $N\uparrow$.

2. التوزيع التكراري المتجمع النازل

التكرار المجتمع النازل

الفئات	التكرارات F_i		الحد الأدنى
10−19	4	40	10 فأكثر
20−29	9	36	20 فأكثر
30−39	15	27	30 فأكثر
40−49	8	12	40 فأكثر
50−59	4	4	50 فأكثر
	40		

التمرين الثاني:

1. تكوين التوزيع التكراري

درجة الطالب في الإختبار كمي مستمر ولكي يتم تبويبها تتبع الخطوات التالية:

- حساب المدى: R

$$R = X_{mxx} - X_{min}$$
$$R = 94 - 55 = 39$$

2. تحديد عدد الفئات:

$$K = 1 + 3.322\log(N)$$
$$K = 1 + 3.322\log(70)$$
$$K \cong 8$$

3. تحديد طول الفئة:

$$L = \frac{R}{K} = \frac{39}{8} = 5.47$$
$$L \cong 6$$

إذن طول الفئات المناسب لإفراغ هذه البيانات في جدول تكراري مستمر متصل هو 5 أمّا عدد الفئات المناسب فهو 8.

وبالتالي فالجدول المطلوب هو

الفئات	التكرار المطلق n_i	التكرار النسبي $F_i = \frac{n_i}{\sum n_i}$	التكرار النسبي المئوي%
]55−61]	10	$F_i = \frac{10}{70} = 0.1428$	%14.28
]61−67]	12	$\frac{12}{70} = 0.1714$	%17.14
]67−73]	13	$\frac{13}{70} = 0.1857$	%18.57
]73−79]	16	$\frac{16}{70} = 0.2286$	%22.86
]79−85]	10	$\frac{10}{70} = 0.1428$	%14.28
]85−91]	4	$\frac{4}{70} = 0.0571$	%5.71
]91−97]	5	$\frac{5}{70} = 0.0714$	%7.14
المجموع	70	1	100

4. نسبة الطلاب الحاصلين على درجات ما بين 70 وأقل من 80 المئوي هو مجموع التكرارين النسبي للفئتين 4 و5 أي 14.28+22.86 ويمكن التقريب أي كتابتها 37.2%=14.3+22.9 أي 37.2% تحصلوا على درجات بين 70 و80.

5. نسبة الطلاب الحاصلين على درجات أقل من 70 هو مجموع التكرارات للفئات 1 و2 و3.

(نكتبها مقربة عشريا مباشرة) أي: 50%=18.6+17.14+14.3.

هناك حوالي 50" من الطلاب تحصلوا على أقل من 70 درجة.

6. نسبة الطلاب الحاصلين على درجة 80 وأكثر، هو مجموع التكرار النسبي المئوي للفئات 6-7-8.

أي 12.8%=2.9+4.3+5.71

أي 12.8% تحصلوا على 80 درجة فأكثر.