جامعة محمد خيضر بسكرة
كلية العلوم الاقتصادية والتجارية
وعلوم التسيير

مجال العلوم الاقتصادية والتسيير
والعلوم التجارية *LMD-SEGC*
السنة الأولى ليسانس

محاضرات في مقياس

الإحصـــــاء الوصفي

المحور الثاني: عرض المعطيات:

الجزء الأول: العرض الجدولي للمعطيات.

إعداد الدكتور:

الهاشمي عبابسة.

h.ababsa@univ-biskra.dz

المحور الثاني العرض الجدولي للمعطيات.

ينتمي عرض المعطيات إلى المرحلة الثانية من مراحل المنهج الإحصائي كما ذكرنا سابقا، ويتم بكيفيتين: الكيفية الجدولية، والكيفية البيانية.

I. العرض الجدولي للمعطيات: (تبويب البيانات)

يتمثل في وضع المعطيات الخام المحصل عليها في شكل جداول إحصائية تسمى "**التوزيعات التكرارية**"، وذلك بعد تنظيم هذه المعطيات وترتيبها.

أ- التوزيعات التكرارية :

التوزيع التكراري هو جدول إحصائي لُخصت فيه المعطيات الخام في طوائف أو فئات، مع تحديد عدد العناصر المنتمية لكل فئة، وهو ما يسمى "**بتكرار الفئة**". وفيما يأتي أمثلة عن توزيعات تكرارية لمتغيرات كمية ونوعية.

- ✓ ***المثال الأول: لمتغير كمي متقطع.***

نفرض أننا رمينا زهرة نرد 10 مرات، فحصلنا على النتائج التالية: 1، 1، 2، 2، 2، 3، 4، 5، 5، 6 تمثل هذه السلسلة العددية معطيات مرتبة ترتيبا تصاعديا، يمكن تلخيصها في توزيع تكراري كما هو مبين في الجدول رقم (2-1) .

الأستاذ / هاشمي عبابسة

الجدول رقم (2-1): نتائج رمي زهرة نرد 10 مرات.

قيم المتغير X	1	2	3	4	5	6	المجموع
التكرار n	2	3	1	1	2	1	**10**

المصدر: معطيات المثال الأول.

- ✓ ***المثال الثاني: لمتغير كمي مستمر.***

لنفرض أننا حصلنا على أوزان 100 طالب من احدى الجامعات، وبعد ترتيبها لخصت في التوزيع التكراري الآتي: (انظر الجدول رقم 2-2).

الجدول رقم (2-2): أوزان 100 طالب

الفئات (كغ)	التكرارات (n_i)
62 – 60	5
65 – 63	18
68 – 66	42
71 – 69	27
74 – 72	8
المجموع	**100**

المصدر: افتراضي.

✓ ***المثال الثالث: لمتغير نوعي.*** ***الجدول رقم (2-3): عدد المواليد حسب الجنس***

وليكن الجنس، حيث سجلنا عدد المواليد من الذكور والإناث في إحدى المستشفيات خلال عام 2005 ولخصناها في التوزيع التكراري الآتي: (انظر الجدول رقم 2-3).

صفة المتغير	التكرارات (n_i)
ذكور	2055
إناث	2180
المجموع	**4235**

المصدر: افتراضي.

الأستاذ / هاشمي عبابسة

ب- الفئات:

الفئة عبارة عن فضاء أو إطار يضم طائفة من المشاهدات المتقاربة أو المتطابقة في القيم وفق معيار معين، أو المشتركة في الصفة نفسها.

مثلا تمثل الفئة الأولى في الجدول رقم (2-2) أعلاه عدد الطلبة الذين تتراوح أوزانهم بين 60 كيلوغراما إلى 62 كيلوغراما وهم خمسة طلبة، وهذا ما يطلق عليه تكرار الفئة الأولى، ويرمز له بالرمز n_1. $(n_1 = 5)$

وفيما يلي بعض المفاهيم المتعلقة بالفئات:

1. الحدود العادية (الظاهرية) والحدود الفعلية (الحقيقية) للفئات:

يسمى الرقمان 60 و 62 على الترتيب الحدين العاديين للفئة الأولى من الجدول رقم (2-2)، حيث يدعى الأول حدّها الأدنى ويسمى الآخر حدها الأعلى.

لكن هذه الفئة تضم -نظريا- كافة الأوزان من 59.5 كغ إلى 62.5 كغ، يسمى هذان الرقمان **الحدين الفعليين** الأدنى والأعلى للفئة الأولى على الترتيب.

ولتحديد الحد الأعلى الفعلي لفئة ما نجمع حدها الأعلى مع الحد الأدنى للفئة الموالية ونقسم المجموع على 2 فنحصل على الحد الأعلى الفعلي لهذه الفئة الذي يمثل -في الوقت نفسه- الحد الأدنى الفعلي للفئة الموالية، نرمز للحدين الفعليين الأعلى والأدنى لفئة ما بالرمزين B_{min} و B_{max} على الترتيب. والجدول الآتي يبين الحدود الفعلية للفئات الظاهرة في الجدول رقم (2-2) السابق:

الجدول رقم (2-4): الحدود الفعلية لأوزان 100 طالب

الفئات (كغ)	الحدود الفعلية	التكرارات (n_i)
60 – 62	59.5 – 62.5	5
63 – 65	62.5 – 65.5	18
66 – 68	65.5 – 68.5	42
69 – 71	68.5 – 71.5	27
72 – 74	71.5 – 74.5	8
المجموع		**100**

المصدر: الجدول رقم (2-2)

ما الفرق بين الحدود العادية والحدود الفعلية للفئات؟

هناك فارقان جوهريان بين الحدود العادية والحدود الفعلية:

- ✓ *الفعلية متصلة، حيث تبدأ الفئةُ من حيث انتهت سابقتُها، بينما العادية منفصلة، بينها فراغ ثابت، يساوي هذا الفراغ مقدار وحدة الدقة.*
- ✓ *الفعلية دائما تزيد عن العادية بدرجةٍ في الدقة: فإذا كانت العادية صحيحةً فإن الفعلية تكون برقم بعد الفاصلة، وإذا كانت العادية برقم بعد الفاصلة تكون الفعلية برقمين بعد الفاصلة وهكذا...*

لكن ما هي وحدة الدقة؟

لتبسيط هذا المفهوم، نقول إن وحدة الدقة هي أصغر انتقال رقمي ضمن سلسلة من الأعداد. فمثلا إذا كان لدينا سلسلة أرقام صحيحة، فإن وحدة الدقة=1 ، وإذا كانت هذه السلسلة ذات رقم واحد بعد الفاصلة، فإن وحدة الدقة = 0.1 ، وإذا كانت هذه السلسلة ذات رقمين بعد الفاصلة، فإن وحدة الدقة = 0.01 وهكذا...

2. الفئة المفتوحة:

قد نجد في بعض الحالات فئة ليس لها أحد حديها مثلا 65 سنة فأكثر أو أقل من 15 سنة، يسمى هذا النوع من الفئات "فئات مفتوحة"، وتقعان في أول التوزيع أو في آخره، والذي يسمى بدوره "توزيعا مفتوحا".

3. طول الفئة (حجم الفئة):

يساوي الفرق بين حدّيها "الفعليين الأعلى والأدنى"، فمثلا طول الفئة الأولى هو 62.5 - 59.5 = 3 كغ.

تنبيه: يمكن استخدام الحدود العادية لحساب طول الفئة، لكن بتعديل بسيط على القانون السابق، فيكون طول الفئة يساوي "**الفرق بين حديها العاديين الأعلى والأدنى زائد وحدة الدقة**".

4. مركز الفئة:

مركز الفئة هو منتصفها، يرمز له بالرمز C. وتبرز أهميته في أنه القيمة التي تمثل بقية المشاهدات المنتمية إلى الفئة، بعبارة أخرى، يساعد مركز الفئة على التعامل رقميا وحسابيا مع هذه الفئة كما سنوضح لاحقا.

يحسب مركز الفئة باستخدام الحدود الفعلية أو الحدود العادية للفئات كما يلي:

$$C = \frac{B_{min} + B_{max}}{2}$$

مثلا مركز الفئة الأولى

$$C_1 = \frac{60 + 62}{2} = \frac{59.5 + 62.5}{2} = 61kg$$

ج- التوزيع التكراري المتجمع:

ويسمى أيضا التوزيع التكراري التراكمي، وينتج عن تجميع أو مراكمة التكرارات صعودا أو نزولا، فهو نوعان:

1. **التوزيع التكراري المتجمع الصاعد**: هو جدول إحصائي يضم التكرارات المتجمعة الصاعدة للفئات، يمكن الحصول على التكرار المتجمع الصاعد لأية فئة بحساب مجموع التكرارات لجميع الفئات التي تسبقها بما في ذلك تكرار هذه الفئة، وبعبارة أخرى يساوي مجموع القيم "الأقل من حدها الأعلى الفعلي".

فمثلا التكرار المتجمع الصاعد للفئة الثالثة من الجدول رقم (2-2) هو 5+18+42= 65 طالبا.

أي أن هناك 65 طالبا أوزانهم أقل من 68.5 كغ.

يسمى الجدول الذي يضم التكرارات المتجمعة الصاعدة لكل الفئات التوزيع التكراري المتجمع الصاعد، أو التوزيع على أساس " **الأقل من...**" كما جاء في بعض المراجع (انظر الجدول رقم 2-5 أدناه).

يرمز عادة للتكرار المتجمع الصاعد للفئة i بالرمز $F_i\uparrow$

2. التوزيع التكراري المتجمع النازل:

هو جدول إحصائي يضم التكرارات المتجمعة النازلة للفئات، يمكن الحصول على التكرار المتجمع النازل لأية فئة بحساب عدد المشاهدات التي تساوي حدها الأدنى الفعلي أو أكثر، أي يساوي تكرارها زائد تكرارات الفئات التي بعدها. ويمكن حسابه بعدة طرق:

مثلا في الجدول رقم (2-5) أدناه:

التكرار المتجمع النازل للفئة الثالثة = 42+27+8=77 طالبا.

= التكرار المتجمع النازل للفئة 02 - التكرار المطلق للفئة 02

= 95 -18 =77 طالبا.

= مجموع التكرارات - التكرار المتجمع الصاعد للفئة 02.

= 100 - 23 = 77 طالبا.

الأستاذ/ هاشمي عبابسة

يسمى الجدول الذي يضم التكرارات المتجمعة النازلة للفئات التوزيع المتجمع النازل أو التوزيع على أساس "**أكثر من....**"، ويرمز له عادة بالرمز $F_i\downarrow$ (انظر الجدول رقم 2-5 أدناه).

د- التوزيع التكراري النسبي:

هو جدول إحصائي يضم التكرارات النسبية للفئات. يمكن حساب التكرار النسبي f_i لأية فئة بقسمة تكرارها المطلق n_i على مجموع التكرارات n ويمكن التعبير عنه بنسبة مئوية.

مثلا: من الجدول رقم (2-4) أعلاه: $f_3 = \frac{42}{100} = 0.42 = 42\%$

ولهذا فإن مجموع التكرارات النسبية يجب أن يساوي 1 أو 100٪ .

يسمى الجدول الذي نحصل عليه في النهاية "التوزيع التكراري النسبي". (أنظر الجدول رقم 2-5 أدناه).

الجدول رقم (2-5): التكرارات المتجمعة والنسبية.

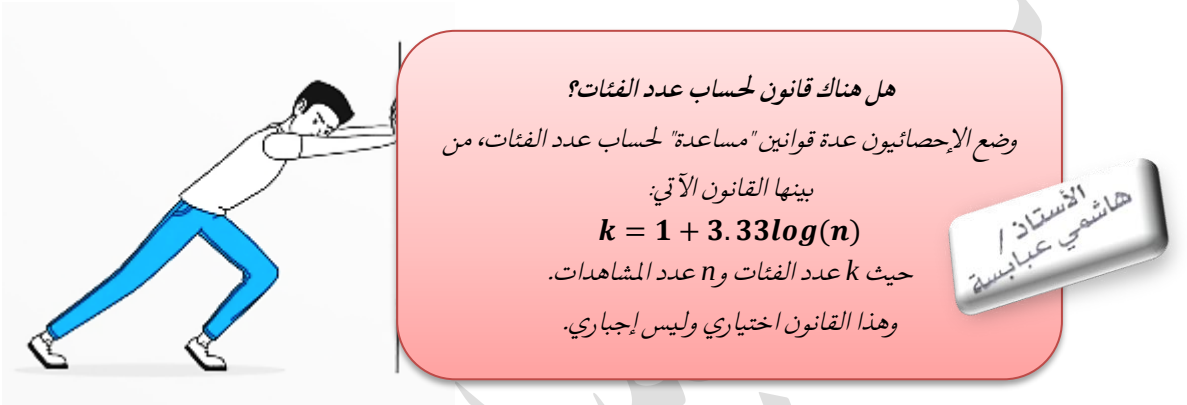
الفئات (كغ)	الحدود الفعلية	n_i	التكرار المتجمع الصاعد		التكرار المتجمع النازل		f_i	f_i%
			الحدود العليا الفعلية	$F_i\uparrow$	الحدود الدنيا الفعلية	$F_i\downarrow$		
			أقل من 59.5	0	أكثر من 59.5	100		
62 – 60	62.5 – 59.5	5	أقل من 62.5	5	أكثر من 62.5	95	0.05	5
65 – 63	65.5 – 62.5	18	أقل من 65.5	23	أكثر من 65.5	77	0.18	18
68 – 66	68.5 – 65.5	42	أقل من 68.5	65	أكثر من 68.5	35	0.42	42
71 – 69	71.5 – 68.5	27	أقل من 71.5	92	أكثر من 71.5	8	0.27	27
74 – 72	74.5 – 71.5	8	أقل من 74.5	100	أكثر من 74.5	0	0.08	8
المجموع		**100**		/		/	**1.00**	**100**

المصدر: الجدول رقم (2-4) أعلاه.

ه- **خطوات تكوين توزيع تكراري:**

لإفراغ (تبويب) مجموعة من المعطيات الخام ضمن توزيع تكراري نتبع الخطوات التالية:

1. **حساب مدى البيانات**: والذي يساوي الفرق بين أكبر قيمة وأصغر قيمة بين قيم المعطيات.
2. **تحديد عدد ملائم من الفئات**: وقد اختلف المؤلفون حول العدد الملائم من الفئات؛ فمنهم من يرى أن العدد المناسب بين 4 و 15 فئة، ومنهم من يرى أنه بين 5 و 25 فئة، وفريق ثالث يرى أن العدد بين 4 و 24 فئة... وعلى أية حال فإن العدد خاضع في النهاية إلى ذوق الباحث وحاجته وهدفه من البحث.



3. **تحديد طول الفئات**: يتم ذلك بقسمة المدى على عدد الفئات المحدد في الخطوة السابقة، فنحصل على نتيجة تساعدنا في تحديد طول ملائم للفئات.
4. **تحديد الحدود العادية للفئات**: حيث يجب أن يكون الحد الأدنى للفئة الأولى يساوي أقل قيمة في المعطيات أو أصغر قليلا، وأن يكون الحد الأعلى للفئة الأخيرة أكبر من أو يساوي أعلى قيمة في المعطيات. (يُستحسن اتخاذ أصغر قيمة في المعطيات كحد أدنى للفئة الأولى).
5. **حساب الحدود الفعلية للفئات**: وفقا للقانون السابق الذكر، وهي خطوة غير إجبارية.
6. **تحديد عدد المشاهدات في كل فئة**: وهو ما يسمى بتكرار الفئة n_i.

مثال: يمثل الجدول رقم (2-6) أسفله نتائج اختبار الذكاء لمئة طالب في إحدى الكليات.

الجدول رقم (2-6): اختبار الذكاء لمئة طالب في إحدى الكليات.

81	**134**	89	105	96	82	84	111	102	109
102	117	107	100	91	79	95	90	107	113
116	93	103	75	82	121	93	90	109	80
104	82	104	96	73	92	78	90	86	99
108	83	94	81	76	80	108	104	103	101
127	110	80	72	121	109	87	93	72	104
90	81	119	111	106	109	103	91	105	111
100	77	103	101	110	92	89	83	93	122
117	**67**	114	91	117	93	79	103	99	99
119	100	89	84	98	84	82	93	93	108

*المصدر: عبد الجبار توفيق البياتي (2008): **الإحصاء وتطبيقاته في العلوم النفسية والاجتماعية**، الطبعة الأولى، إثراء للنشر والتوزيع، عمان، ص 50.*

المطلوب: تبويب هذه البيانات في توزيع تكراري.

للقيام بذلك نتبع الخطوات التالية:

- ✓ **حساب مدى البيانات**: ويساوي أعلى قيمة – أدنى قيمة = 134 – 67 = 67.
- ✓ **تحديد عدد ملائم من الفئات**: كما ذكرنا في الدرس، اختلف الإحصائيون في العدد الملائم من الفئات؛ فمنهم من يفضل استخدام ما بين 10 و 20 فئة، ومنهم من يفضل ما بين 7 و 15 فئة، وذهب آخرون الى توسيع المجال ليكون ما بين 5 و 25 فئة.

وعموما ليس هناك قاعدة واحدة لتحديد عدد ملائم من الفئات، إذ إن الأمر يتوقف على عدد البيانات الخام وكذلك على ذوق الباحث وهدفه من الدراسة؛ فإذا كان لدينا 30 مشاهدة فإنه من التضليل تصنيفها في 10 فئات، كما أنه من غير المقبول تبويب 3000 مشاهدة في خمس فئات.

وقد وضع الاحصائيون عدة قوانين مختلفة تساعد الباحث في تحديد عدد ملائم من الفئات، من ذلك القانون الآتي: عدد الفئات = 1 + 3.33 لغ n = 1 + 3.33 لغ 100 = 7.67 = 8.

اذن نختار **ثمان فئات**.

- ✓ **حساب أطوال الفئات** (مع تحديد الحدود العادية والفعلية للفئات): سنضع فئات متساوية الطول، طول كل واحدة منها يساوي (المدى / عدد الفئات) = (67 / 8) = 8.375 ≅ 9. (سنجرب 9 كطول للفئة). ونعتبر أقل قيمة في البيانات الخام (67) حدا أدنى للفئة الأولى. (وحدة الدقة تساوي 1 لأن البيانات الخام عبارة عن ارقام صحيحة كما أوضحنا سابقا)

الجدول رقم (2-7): تجريب 9 كطول لفئات درجات الذكاء.

الفئات (حدود عادية)	الفئات (حدود فعلية)
75 – 67	75.5 – 66.5
84 – 76	84.5 – 75.5
93 – 85	93.5 – 84.5
102 – 94	102.5 – 93.5
111 – 103	111.5 – 102.5
120 – 112	120.5 – 111.5
129 – 121	129.5 – 120.5
138 – 130	138.5 – 129.5

المصدر: بيانات الجدول رقم (2-6) أعلاه.

الأستاذ / هاشمي عبابسة

- ✓ **تحديد التكرارات**: وذلك اعتمادا على ما يسمى بتقنية "الحُزم"، حيث نحدد عدد المشاهدات المنتمية لكل فئة. حيث كل حزمة تمثل 5 مشاهدات (يمكن استخدام المربع ذي القطر الواحد)

شكل الحزمة: 𝍸 شكل المربع ذي القطر الواحد: ⧅

كل هذه الخطوات ملخصة في التوزيع التكراري الآتي: (أنظر الجدول رقم 2-8 أسفله)

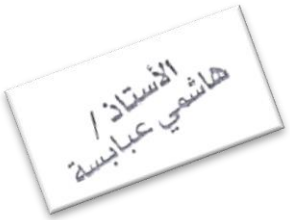
الجدول رقم (2-8): التوزيع التكراري لدرجات ذكاء عينة ن الطلبة.

الفئات (حدود عادية)	الفئات (حدود فعلية)	تقنية الحزم	التكرارات n_i
75 – 67	75.5 – 66.5	𝍸	5
84 – 76	84.5 – 75.5	𝍸 𝍸 𝍸 𝍸	20
93 – 85	93.5 – 84.5	𝍸 𝍸 𝍸 𝍸 I	21
102 – 94	102.5 – 93.5	𝍸 𝍸 𝍸	15
111 – 103	111.5 – 102.5	𝍸 𝍸 𝍸 𝍸 𝍸 I	26
120 – 112	120.5 – 111.5	𝍸 III	8
129 – 121	129.5 – 120.5	IIII	4
138 – 130	138.5 – 129.5	I	1
100	المجموع		

المصدر: بيانات الجدولين (2-6) و (2-7) أعلاه.

تنبيه هام.

عادة ما يتم حذف العمود الخاص بالحزم لأنه مجرد عمود مساع لتحديد التكرارات، لكن أثناء إنجاز هذا العمل ضمن امتحان لا بد من ترك هذا العمود ليفهم الأستاذ المصحح الجهد الذي بذله الطالب لتحديد التكرارات.



و- العرض الجدولي في حالة خاصيتين أو متغيرين (التوزيع التكراري المزدوج).

يستخدم الجدول المزدوج عند دراسة خاصيتين أو متغيرين في الوقت نفسه للمجتمع الإحصائي ذاته، يرمز للخاصية الأولى بالرمز X وقيمها x_i ، و يرمز الخاصية الثانية بالرمز Y وقيمها y_i .

وفي هذا الإطار يمكن أن نميز ثلاث حالات:

- ✓ الحالة الأولى: المتغيران نوعيان.
- ✓ الحالة الثانية: أحدهما كمي والآخر نوعي.
- ✓ الحالة الثالثة: المتغيران كميان.

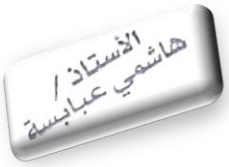
وعموما يمكن تمثيل الجدول المزدوج كما يأتي: (أين k يمثل عدد الأسطر و l يمثل عدد الأعمدة).

الجدول رقم (2-9): توزيع تكراري لخاصيتين أو متغيرين X و Y.

Y_j / X_i	y_1	y_2	…….	y_l	المجموع
x_1	n_{11}	n_{12}	…….	n_{1l}	$n_{1.}$
x_2	n_{21}	n_{22}	…….	n_{2l}	$n_{2.}$
.	.	.	.	.	.
.	.	.	…….	.	.
.	.	.	.	.	.
x_k	n_{k1}	n_{k2}	…….	n_{kl}	$n_{k.}$
المجموع	$n_{.1}$	$n_{.2}$	…….	$n_{.l}$	$n_{..}= n$

المصدر: افتراضي.

- تسمى التكرارات «$n_{i.}$ »المجموع الحدي أو التكرارات الحدية للمتغير X.
- تسمى التكرارات «$n_{.j}$ »المجموع الحدي أو التكرارات الحدية للمتغير Y.



يمكن كتابة هذه المجاميع الحدية كما يلي:

✓ **بالنسبة للمجاميع الحدية للأسطر**:

$$n_{i.} = n_{i1} + n_{i2} + \cdots + n_{il} = \sum_{j=1}^{l} n_{ij}$$

مثلا: لحساب المجموع الحدي للسطر الأول (أي $n_{1.}$) ، نضع i تساوي 1 في القانون السابق ونجمع جميع خانات الأعمدة والواقعة في السطر الأول. وهكذا مع السطر الثاني والثالث ...

✓ **بالنسبة للمجاميع الحدية للأعمدة**:

$$n_{.j} = n_{1j} + n_{2j} + \cdots + n_{kj} = \sum_{i=1}^{k} n_{ij}$$

مثلا: لحساب المجموع الحدي للعمود الأول (أي $n_{.1}$) ، نضع j تساوي 1 في القانون السابق ونجمع جميع خانات السطر والواقعة في العمود الأول. وهكذا مع العمود الثاني والثالث ...

✓ **بالنسبة للمجموع الكلي**: نحصل عليه بجمع المجاميع الحدية للأسطر أو للأعمدة كما يلي:

$$n = \sum_{i=1}^{k} n_{i.} = \sum_{i=1}^{k} \left(\sum_{j=1}^{l} n_{ij} \right) = n..$$

$$n = \sum_{j=1}^{l} n_{.j} = \sum_{j=1}^{j} \left(\sum_{i=1}^{k} n_{ij} \right) = n..$$

أي أن مجموع قيم السطر الأخير تساوي مجموع قيم العمود الأخير وتساوي جميعا مجموع التكرارات الكلي.

الأستاذ / هاشمي عبابسة

أمثلة:

- **عن الحالة الأولى: المتغيران نوعيان.**

لنفرض أننا سحبنا عينه مكونة من 130 طالبا ودرسنا خاصيتي لون العينين X ولون الشعر Y ، ثم أنشأنا التوزيع التكراري التالي (انظر الجدول رقم 2-10).

الجدول رقم (2-10): لون الشعر ولون العينين لعينة من الطلبة.

Y / X	الأسود	الأشقر	البني	الأصفر	المجموع
الأسود	18	7	16	3	**44**
الأزرق	5	20	5	3	**33**
الأخضر	2	12	3	3	**20**
البني	10	6	12	5	**33**
المجموع	**35**	**45**	**36**	**14**	**130**

المصدر: افتراضي.

- **عن الحالة الثانية: متغير كمي والآخر نوعي.**

لنفرض أننا درسنا نقاط 30 طالبا في أحد المقاييس (X) بين ذكور وإناث (Y).

الجدول رقم (2-11): نقاط 20 طالبا في أحد المقاييس.

X / Y	20 – 29	30 – 39	40 – 49	50 – 59	60 – 69	70 – 79	المجموع
ذكـــور	1	1	4	4	2	1	**13**
إنـــاث	1	2	5	8	1	0	**17**
المجموع	**2**	**3**	**9**	**12**	**3**	**0**	**20**

المصدر: افتراضي.

- **عن الحالة الثالثة: المتغيران كميان.**

مثلا درسنا الوزن بالكيلوغرام (X) والطول بالسنتيمتر (Y) لمئة طالب جامعي.

الجدول رقم (2-12): الوزن والطول لمئة طالب.

Y / X	150 – 159	160 – 169	170 – 179	180 – 189	190 – 199	التكرارات (n_i)
60 – 62	0	1	2	1	1	**5**
63 – 65	1	2	10	3	2	**18**
66 – 68	3	5	21	9	4	**42**
69 – 71	1	3	15	5	3	**27**
72 – 74	0	2	3	2	1	**8**
المجموع	**5**	**13**	**51**	**20**	**11**	**100**

المصدر: افتراضي.

انتهى الجزء الأول من المحور الثاني: العرض الجدولي للمعطيات.

أستاذ المقياس الدكتور الهاشمي عبابسة.
h.ababsa@univ-biskra.dz