

محاضرات مادة الاحصاء الهندسي

المرحلة الاولى للعام الدراسي 2013 – 2014 م. ناريمان يحيى

**المحاضرة الثانية**

**تنظيم البيانات وعرضها جدولياً**

**1-2 تنظيم البيانات وتلخيصها وعرضها جدولياً**

بعد جمع البيانات فإنها تكون بيانات أولية غير منظمة عددياً وتصبح دراستها او استنتاج أي شيء منها. ولذلك دعت الحاجة الى تلخيص هذه البيانات بصورة يسهل فهمها و استنتاج بعض النتائج الاولية منها ولتوضيح ذلك نعتبر المثال الاتي:

اذا كان لديك تقديرات 60 طالبا في مادة الاحصاء مثلا وكالتالي:

D	B	E	C	D	B	D	C	E	A
B	E	C	D	B	D	D	A	E	C
C	D	A	C	E	D	C	C	D	B
D	E	D	D	A	D	D	C	D	C
D	A	B	D	B	D	C	D	C	E
D	B	C	C	E	D	C	C	D	A

البيانات السابقة بوضعها الحالي تجعل من الصعب التعرف على الطلاب الحاصلين على تقدير مشترك مثل ممتاز (A) او جيد جداً (B) .... ومن هنا اصبحت الحاجة الى وضع التقديرات وتلخيصها في جدول يسهل دراسته يسمى **جدول التوزيع التكراري** وقد تكون البيانات رقمية مثل درجات الطلاب , او اطوال الطلاب , او قوة المقاومة للمكعبات الخرسانية , او ابعاد الطابوق . ونوضح ذلك بالمثال التالي:

مثال (2-2):

البيانات التالية تمثل درجات 50 طالباً في مادة الاحصاء:

51	95	70	74	73	90	71	74	90	67
91	72	83	89	50	80	72	84	85	69
62	82	87	76	91	76	87	75	78	79
71	96	81	88	64	82	73	57	86	70
80	81	75	85	74	90	83	66	77	91

البيانات السابقة بوضعها الحالي يصعب دراستها او استنتاج بعض المؤشرات منها. فمثلاً ما هو عدد الطلبة الذين حصلوا على 70 درجة فأكثر؟ او عدد الطلبة الذين حصلوا على درجات تتراوح بين 80 و90 درجة... الخ؟ ولذلك فإن اول مرحلة للتحليل الإحصائي تتكون من تصميم جدول التوزيع التكراري.

وتنظم وتلخص البيانات الاحصائية فيما يسمى **بالتوزيع التكراري (الجدول)** **Frequency Distribution** وهو عبارة عن جدول يلخص البيانات الاولية فيوزعها على **فئات او أصناف Classes** ويحدد عدد الافراد او الأشياء التي تنتمي الى كل فئة أو صنف, ويسمى هذا العدد **تكرار الفئة frequency**, وعادة يرمز له (f) ولإتمام ذلك ينبغي ان تفرغ البيانات الاحصائية في جدول يحوي على ثلاث خانات. الخانة الاولى او العمود الاول تكتب فيه الفئات للبيانات وفي الخانة الثانية توضع علامات وهي عبارة عن حزم مكونة من خمسة خطوط, اربعة منها رأسية والخامس مائل يحزم الخطوط الأربعة الرأسية وبذلك تصبح الحزمة على الصورة (###) وفي الخانة الثالثة الاخيرة يكتب مجموع العلامات امام كل فئة على حدة ومجموع هذه العلامات في كل فئة تسمى **التكرار** لهذه الفئة ويكون مجموع التكرارات دائماً مساوياً الى العدد الكلي للبيانات ( $\sum_{i=1}^k f_i = n$ ). وبذلك يكون جدول تفرغ البيانات الاحصائية في مثال (1-2) وهو تقديرات النجاح للطلاب كالتالي:

**جدول (1-2) يبين تفرغ وتوزيع التقديرات للطلاب في مثال (1-2)**

افصالات	املاعات	اركتلرا د(د الاطلب)
A	### /	6
B	### ///	8
C	### ### ### /	16
D	### ### ### ### //	22
E	### ///	8
او مجموع		60

ومن هذا الجدول نكون جدولاً اخر يسمى **الجدول التكراري او جدول التوزيع التكراري** الذي يتكون من خانتين. الاولى تمثل الفئة أو الصنف والثانية تمثل التكرار كما هو مبين بجدول رقم (2-2) كما يلي:

## جدول (2-2)

افصالات	التكرارا
A	6
B	8
C	16
D	22
E	8
اومجماع	60

واحيانا يكتب الجدول السابق في صورة افقية كما يلي:

افصلا	A	B	C	D	E	اومجماع
اركتلرا	6	8	16	22	8	60

والبيانات السابقة تسمى بيانات وصفية تصف تقدير الطالب فتم عمل الجدول التوزيع التكراري لها كالسابق , اما بالنسبة لدرجات الطلاب فتسمى بيانات كمية او عددية والتي تهمننا في دراستنا في مجال الهندسة المدنية , فيتم عمل جدول التكرار لها حسب الخطوات التالية:

- 1- ايجاد المدى الكلي R للتوزيع الذي يمثل الفرق بين اكبر واصغر قيمة للبيانات.
- 2- ايجاد عدد الفئات أو الاصناف حسب صيغة ستور جيس

$$k = 1 + 3.322 \log n$$

حيث ان:

k : عدد الفئات أو الأصناف ويقرب دائما الى اقرب عدد صحيح.

n : العدد الكلي للبيانات.

- 3- ايجاد طول الفئة أو الصنف L حيث ان:

$$L = R / k$$

ويفضل ان تكون اطوال الفئات او الاصناف كلها متساوية لتسهيل الحل ولكن في بعض الاحيان يتم استخدام اطوال اصناف مختلفة وخاصة للصنفين الاول والاخير وذلك في حالة وجود قيم متطرفة.

فبالنسبة للمثال (2-2) سيتم عمل جدول التوزيع التكراري بعد ان يتم ترتيب البيانات تصاعديا لتسهيل الحل , وحسب الخطوات السابقة , كما يلي:

50	51	57	62	64	66	67	69	70	70
71	71	72	72	73	73	74	74	74	75
75	76	76	77	78	79	80	80	81	81
82	82	83	83	84	85	85	86	87	87
88	89	90	90	90	91	91	91	95	96

$$1- R = 96 - 50 = 46$$

$$2- k = 1 + 3.322 * \log 50 = 6.64$$

$$\text{use } k = 7$$

$$3- L = 46 / 7 = 6.57$$

$$\text{use } L = 7$$

4- **ايجاد نهايات الصنف (class limits)** حيث ان لكل صنف نهايتان دنيا وعليا ويتم اختيار نهايات الاصناف بحيث لا يكون هناك مجال للالتباس او الشك في موقع اي قراءة من القراءات ولذلك تترك فجوة بين صنف واخر على ان لا يتعدى عرضها دقة القراءات. وعليه فان مثال (2-2) سيتم حساب نهايات الاصناف له كالتالي:

\*للصنف الاول: النهاية الدنيا ( $lc_1$ ) (lower class) ستمثل اصغر قيمة للبيانات.

\*اما بالنسبة لبقية الاصناف فسيتم حساب النهايات الدنيا كالتالي :

$$lc_i = uc_{i-1} + \text{accuracy}$$

حيث ان :

accuracy: تعني وحدة دقة القياس (acu)

بينما النهايات العليا لكل الاصناف , بضمنها الاول يتم حسابها كما يلي:

$$uc_i = lc_i + L - \text{acu}$$

$$lc_1 = 50$$

$$uc_1 = 50 + 7 - 1 = 56$$

$$lc_2 = 56 + 1 = 57$$

$$uc_2 = 57 + 7 - 1 = 63$$

$$lc_3 = 63 + 1 = 64 \dots \text{etc};$$

فيكون جدول (2-3) كالتالي:

عدد الفئات	الفئات أو الاصناف	العلامات	التكرارات ( $f_i$ )
1	50 – 56	//	2
2	57 – 63	//	2
3	64 – 70	###/	6
4	71 – 77	### ### ////	14
5	78 – 84	### ###/	11
6	85 – 91	### ### ///	13
7	92 – 98	//	2

