الجمهورية الجزائرية الديمقراطية الشعبية

وزارة التعليم العالي والبحث العلمي

جامعة الشهيد حمة لخضر الوادي

قسم العلوم الاجتماعية

محاضرات في مقياس المعالجة الإحصائية للبيانات

محاضرة رقم 02 مراجعة عامة حول الاحصاء الوصفي (مقاييس النزعة المركزية ومقاييس التشتت)

موجهة لطلبة السنة أولى ماستر إرشاد وتوجيه

إعداد الأستـــــــاذ:

د. محمد سبــع

السنة الجامعية: 2020 / 2021

محتوى المادة:

➢ 2. مراجعة عامة حول مبادئ الإحصاء الوصفي

1.2. مقاييس النزعة المركزية

نسعى من وراء دراستنا لظاهرة إحصائية معينة إلى استنتاج صفة مميزة ما أو أكثر من صفة فعند دراستنا لدرجات طلاب السنة التمهيدية في مادة الرياضيات نستنتج أن معظم الدرجات تميل إلى التمركز حول نقطة أو درجة معينة وهي درجة النجاح بينما نجد أن الدرجات المتطرفة الأخرى من الدرجات المدروسة ، قليلة نوعاً ما وهذه الدرجات ، المتطرفة ، هي الدرجات العالية جداً أو المنخفضة جداً أو كلاهما . نسمي هذا الميل بالتمركز أو بالنزعة المركزية ، ونسمي القيم التي تتمركز حولها القيم الأخرى مقاييس النزعة المركزية Measures of Central location ولهذه المقاييس أهمية كبيرة في علم الإحصاء فهي تعطينا فكرة عامة عن قيم الظاهرة المدروسة وبالتالي مقارنة مجموعتين أو أكثر . هناك عدة مقاييس لهذا الغرض منها الوسط الحسابي أو المتوسط (المعدل) Mean الوسيط Median والمنوال Mode .

أولا: الوسط الحسابي أو المتوسط: Mean :

يمثل الوسـط الحسـابي أو المتوسـط (Arithmetic Average or Mean) مقياس النزعة المركزية الأكثر شــهرة والأكثر أهمية في المقاييس المختلفة. وتمثل قيمة الوســـط الحســـابي القيمة التي تتمركز حولها جميع القيم المختلفة للمتغير الكمي. يمكن الحصول على القيمة الحقيقية لمتوسط متغير عشوائي في مجتمع محدود إذا تم التعامل مع كافة القيم في المجتمع. في هذه الحالة يرمز لقيمة الوسـط الحسـابي المحصـل بالرمز μ والتي تمثل معلمة المجتمع. وبافتراض التعامل مع متغير عشوائي X لمجتمع محدود حجمه N فانه يمكن حساب قيمة الوسط الحسابي من خلال الدالة التالية

$$\mu_X = \frac{\sum_{i=1}^{N} X_i}{N} \qquad [1]$$

أمـا في حـال التعامـل مـع عينـة عشـوائية ممثلـة لمجتمـع الدراسـة حجمهـا n قـراءة، فـان معادلـة تقـدير قيمـة الوسـط الحسـابي الـذي يعبر عنـه بـالرمز $\bar{x}$ ويقـرأ (xbar) لمجموعـة مـن القـيم $x_1, x_2, ... x_n$ هـو حاصل قسمة مجموع القيم في العينة على عددها

$$\bar{x} = \frac{x_1 + x_2 + ... + x_n}{n} \qquad [2]$$

ويشار إليها في كتب الإحصاء المتخصصة بأحد المعادليتين التالييتين:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad [3]$$

$$\bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \qquad [4]$$

مثال (01):

تمثل الدرجات التالية درجات مجموعة من الطلبة في مقياس الإحصاء:

5 ، 8 ، 12 ، 10 ، 15 ، 9 ، 11 ، 10

فإن $10+11+9+15+10+12+8+5 = \sum_{i=1}^{n} x_i$

فإن المتوسط الحسابي وفق المعادلة **[3]**: $\bar{x} = \frac{1}{10}[10+11+9+15+10+12+8+5] = \frac{80}{10} = 8$

وهي ذات القيمة باستخدام المعادلة **[4]** : $\bar{x} = \frac{[10+11+9+15+10+12+8+5]}{10} = \frac{80}{10} = 8$

أما إذا كانت البيانات معطاة بجدول توزيع تكراري ذو قيم $x_1, x_2, \ldots x_n$ وتكراراتها على التوالي هي $f_1.f_2.f_3\ldots f_K$:

$$\boxed{\bar{x} = \frac{1}{n}\sum_{i=1}^{k} x_i f_i} \qquad [5]$$

مثال (02): تمثل القيم التالية تكرارات درجات 25 طالبا: 2 ، 3، 4، 5، 6، 7، 8،

تكراراتهم على التوالي كانت 7، 10، 10، 4، 5، 3، 4، 7

درجات الطلبة x_i	التكرار f_i	الدرجة xالتكرار $x_i f_i$
2	**1**	**2**
3	**2**	**6**
4	**3**	**12**
5	**4**	**20**
6	**5**	**30**
7	**4**	**28**
8	**2**	**16**
9	**4**	**36**
المجموع	**25**	**150**

$$\bar{x} = \frac{\sum_{i}^{k} x_i f_i}{\sum_{i}^{k} f_i} = \frac{\sum_{1}^{8} x_i f_i}{\sum_{1}^{8} f_i} = \frac{1}{25}[2(1)+3(2)+4(3)+5(4)+6(5)+7(4)+8(2)+9(4)] = \frac{150}{25} = 6$$

مثال (03): باستخدام المعادلة نفسها يمكن حساب المتوسط الحسابي لبيانات مبوبة في شكل فئات انطلاقا من المثال التالي:

فئات درجات قلق الامتحان	التكرار f_i	مراكز الفئات x_i	مركز الفئة x التكرار $x_i f_i$	خطوات الحل:
24–20	6	22	132	1. نحسب مراكز الفئات من خلال حاصل قسمة الحد الأعلى للفئة + الحد الأدنى للفئة /2
29–25	4	27	108	2. نخصص عمود في الجدول لحساب حاصل ضرب مركز الفئة x تكرار الفئةمخ من خلال : $\sum f \cdot x$
34–30	9	32	288	
39–35	8	37	296	3. نحسب مجموع التكرارات من خلال $\sum f$.
44–40	6	42	252	
49–45	7	47	329	4. نعوض النواتج في قانون حساب المتوسط
54–50	5	52	260	
المجموع	45		1665	$\bar{x} = \frac{\sum_i^k x_i f_i}{\sum_i^k f_i} = \frac{1665}{45} = 37$

- الخواص الإحصائية للمتوسط:

– مجموع انحرافات قيم مجموعة بيانات عن وسطها الحسابي دائما يساوي الصفر أي: $\sum_{i=1}^{n}(x_i - \bar{x}) = 0$

– يأخذ جميع القيم في الاعتبار.

– من أكثر المقاييس الوصفية استخداما

– يتأثر بالقيم المتطرفة الكبيرة والصغيرة.

ثانيا: الوسيط Median

يعتبر الوسيط (Median) مقياس آخر للنزعة المركزية، حيث يتم من خلال الوسيط الوصول إلى رقم كمي يمثل القيمة التي تقع في منتصف قيم المتغير الكمي المدروس، لذا فإن الوسيط يمثل القيمة الكمية التي تكون نصف قراءات المتغير الكمي أقل منها بينما النصف الآخر أعلى منها، ولحساب الوسيط لابد أولا من أن يتم ترتيب القيم تصاعديا، حيث يتم ذلك من خلال الترتيب التصاعدي (أو الهابط) العادي في حال البيانات الخام ، والوسيط واحد من القيم الوضعية التي يمكن من خلالها معرفة تموضع الدرجات وعموما يمكن معرفة قيمة الوسيط بعد ترتيب الدرجات تصاعديا أو تنازليا ثم حساب قيمته باستخدام إحدى المعادلتين **[6]** أو **[7]** وذلك بحسب عدد القيم إما أن يكون فرديا أو زوجيا وهي المعادلة التي تسمح لنا بمعرفة رتبة الوسيط وليس الوسيط نفسه.

المعادلة **[6]** إذا كان عدد القيم غير المبوبة فرديا

$$Med = X_{(n+1)/2} \qquad [\,6\,]$$

أما إذا كان عدد القيم زوجيا فإن المعادلة الخاصة بحساب قيمة الوسيط هي:

$$Med = \frac{1}{2}\left(X_{(n/2)} + X_{(n+2)/2}\right) \qquad [\,7\,]$$

مثال (04): إذا كان لدينا مجموعة من القيم عددها فردي هي : 2، 5، 3، 9، 12، 10، 8

1. نقوم بترتيب الدرجات تصاعديا: 2، 3، 5، 8، 9، 10، 12
2. عدد القيم n = 7 $/\ Med = x\,(7+1)/2 = 4$
3. 4x = 8

مثال (05): إذا كان لدينا مجموعة من القيم عددها زوجي هي: 2، 5، 3، 9، 12، 10، 8، 11

1. نقوم بترتيب القيم تصاعديا: 2، 3، 5، 8، 9، 10، 11، 12
2. عدد القيم هنا زوجي باستخدام المعادلة [7] نحصل على:

$$Med = \frac{1}{2}\{x\,(8)/2 + x(8+2)/2\} = \frac{1}{2}(8+9) = 8.5$$

أما إذا كانت البيانات مبوبة في فئات فإن قيمة الوسيط تحسب من خلال المعادلة ويتعين قبل ذلك تحديد الفئة الوسـيطية وهي الفئة التي يزيد أو يســاوي تكرارها المتجمع الصــاعد رتبة الوسـيط $\frac{n}{2}$ حيث n هو مجموع التكرارات .

$$\boxed{Me = L_1 + \left(\frac{\frac{n_i}{2} - \Sigma f_1}{f_{Me}}\right) \times C} \quad \textbf{[8]}$$

حيث:

Me: هو الوسيط

L_1 : الحد الأدنى الفعلي للفئة الوسيطية وهو قيمة الحد الأدنى − 0.5

n_i : مجموع التكرارت $(\sum_{i=1}^{K} f_i)$

$\frac{n_i}{2}$ رتبة الوسيط

Σf_1 التكرار المتجمع الصاعد السابق للفئة الوسيطية

f_{Me} تكرار الفئة الوسيطية

C طول الفئة

مثال (06): انطلاقا من القيم الواردة في الجدول أدناه أوجد قيمة الوسيط

فئات درجات قلق الامتحان	الحدود الحقيقية للفئات	التكرار f_i	مراكز الفئات x_i	التكرار المتجمع الصاعد
24–20	24.5–19.5	6	22	6
29–25	29.5–24.5	4	27	10
34–30	34.5–29.5	9	32	19
39–35	39.5–34.5	8	37	27

44–40	44.5–39.5	6	42	33
49–45	49.5–44.5	7	47	40
54–50	54.5–49.5	5	52	45

من خلال عمود التكرار التجميعي الصاعد نجد أن : $\frac{n}{2}=\frac{45}{2}=22.5$

وبالتالي فالفئة الرابعة تحتوي على الوسـيط وحدها الأدنى 35 وبالتالي فان حدها الادنى الفعلي هو 34.5 فبتطبيق العلاقة السابقة **[8]** نجد أن وسيط البيانات هو

$$\tilde{x}=34.5+\frac{22.5-19}{8}5=36.68$$

ثالثا: المنوال Mode :

يمثل المنوال (Mode) ويشار إليه عادة بالرمز Mod وهو القيمة الأكثر شيوعا من بين القيم المختلفة للمتغير العشوائي محل الدراسة، ويتم تحديد قيمة المنوال من خلال تحديد تكرار جميع القيم المختلفة للمتغير العشوائي محل الدراسة إذا كانت البيانات غير مبوبة (بيانات خام)، بينما يتم الاستعانة بقاعدة رياضية إذا كانت قيم المتغير العشوائي متوفرة في جدول تكراري (بيانات مبوبة).

بالنسبة للبيانات الخام، يتم تحديد قيمة وحيدة للمنوال إذا وجدت قيمة واحدة تكررت أكثر من باقي القيم المختلفة للمتغير العشوائي. كذلك يمكن أن يكون المنوال متمثل بأكثر من قيمة إذا كان هنالك أكثر من قيمة واحدة لها نفس التكرار الأكثر من بين جميع التكرارات المتوفرة. وفي حال عدم تكرر أي قيمة من قيم المتغير العشوائي المختلفة فانه في هذه الحالة لا يكون هنالك منوال بين قيم المتغير العشوائي.

وبما أن المنوال هو القيمة الأكثر تكرارا فإن المنوال من القيم التالية: 1,1,2,3,3,4,5,5,5,6,7 منوالا واحدا وهو $Mod = 5$ ، قد لا يكون لمجموعة من القيم منوال كما هو الحال في القيم 1,2,3,4,5,6,7 وقد يكون لمجموعة من القيم أكثر من منوال واحد كما هو الحال في القيم 1,1,1,2,2,2,3,4,5,5,6 .

و إذا كانت البيانات غير عددية (أي نوعية) كأن نسأل مثلا ستة أشخاص عن أي الألوان المحببة لهم ، قد تكون إجاباتهم كالتالي:

أزرق, أصفر, أبيض, أزرق, أبيض أو أزرق فسيكون المنوال في هذه الحالة هو اللون "الأزرق" .

أما إذا كانت البيانات مبوبة في جدول توزيع تكراري فنحسب المنوال كما يلي : نعين الفئة المنوالية وهي الفئة التي تقابل أكبر قيمة للتكرارات ثم نحسب الفرق بين تكرار الفئة المنوالية وتكرار الفئة التي قبلهاΔ_1 ثم نحسب الفرق بين تكرار الفئة المنوالية وتكرار الفئة التي بعدها وليكن Δ_2 نرمز للحد الأدنى الفعلي للفئة المنوالية بالرمز a ولطول الفئة بالرمز Δ فنحصل على المنوال كما يلي:

$$\boxed{x^{\circ}=a+\frac{\Delta_1}{\Delta_1+\Delta_2}.\Delta} \quad [\,9\,]$$

وكتطبيق مباشر على ذلك ، المنوال للقيم المعطاة بالجدول مثال رقم 06 أعلاه هو

$$. x^{\circ} = 29.5 + \frac{9-4}{5+1} \times 5 = 33.66$$

2.2. مقاييس التشتت (التباين): Measure of Variation

لا تكتمل الإفادة من مقاييس الإحصاء الوصفي دون مقاييس التشتت التي تشكل إضافة إلى مقاييس النزعة المركزة موضوع الشرح والتفصيل في الصفحات السابقة طرفي معادلة الإحصاء الوصفي، إذ غالبا ما تكون مقاييس النزعة المركزية لوحدها غير كافية لتمثل الواقع بشكل كامل أو لمقارنة مجموعتين من المشاهدات أو أكثر، فمن خلال معطيات القيم للمجموعتين التاليتين نقف على عدم كفاية مقاييس التمركز لوحدها كدالة لوصف معالم العينة والمجتمع المستهدف بالدراسة الإحصائية:

مثال (07): المجموعة الأولى : 9، 10، 10، 11، 12، 12، 14، 18

المجموعة الثانية : 2، 8، 9، 11، 12، 16، 19، 19

لوجدنا أن الوسط الحسابي لكل مجموعة هو 12 كما أن الوسيط هو نفسه للمجموعتين ويساوي 11.5 ومع ذلك فهناك فرق واضح بين قيم المجموعتين حيث تختلف مفردات المجموعة الأولى عن مفردات المجموعة الثانية، كما أن قيم المجموعة الثانية موزعة على مدى أوسع من المجموعة الأولى ويمكن أن نقول إن تشتت المجموعة الثانية أكبر منه في المجموعة الأولى، لذلك تمثل مقاييس التشتت الجانب الآخر من المقاييس الإحصائية الأساسية بجانب مقاييس النزعة المركزية، حيث تستخدم تلك المقاييس في وصف البيانات والتعرف على خصائصها، كما تعمل مقاييس التشتت كجزئية مكملة ومهمة جدا بجانب مقاييس النزعة المركزية في عمليات الاستدلال الإحصائي المبنية على عملية التعامل مع البيانات، وينصب الاهتمام عند التعامل مع مقاييس التشتت حول قياس درجة الاختلاف بين القيم المختلفة للمتغير الكمي المدروس، ويتم ذلك من خلال عدة مقاييس مختلفة يهتم كل واحد منها بقياس درجة الاختلاف من زاوية مختلفة، ويمثل التباين والانحراف المعياري بالإضافة إلى المدى مقاييس مختلفة لقياس تشتت المتغيرات الكمية.

يتم الحصول على تصور دقيق عن خصائص المتغير الكمي في حال توفر كل من مقياس النزعة المركزية ومقياس التشتت، حيث تعطي مقاييس النزعة المركزية تصور عن تمركز القيم بينما تعطي مقاييس التشتت تصور عن درجة اختلاف تلك القيم عن بعضها البعض، لذا يمكن القول بأن الاعتماد على مقياس واحد قد لا يغني عن الآخر في عملية الاستدلال الإحصائي، حيث ينتج عنه دوما قصور في المعلومة المعتمد عليها ومن ثم عدم القدرة على قراءة البيانات إحصائيا بشكل سليم.

يتم في الواقع حساب مقاييس التشتت للبيانات الكمية بصيغتيها الخام والمبوبة، ولكن الطريقة المتبعة في عمليات الحساب تختلف باختلاف طبيعة البيانات المدروسة. وبالطبع، كما تم الإشارة إليه

سابقا، تمثل التقديرات المحصلة من البيانات الخام معلومة أدق وأكثر صحة من المعلومة المحصلة من البيانات المبوبة، لذا فانه يجب الاعتماد على البيانات الخام في حال توفرها في عملية حساب مقاييس التشتت، كما يجب قصر الاعتماد على البيانات المبوبة ليتم فقط في حالة عدم توفر الصيغة الخام للبيانات حيث أنها تعطي قيم تقريبية لا ترقى إلى دقة التقديرات المحصلة من خلال استخدام البيانات الغير مبوبة.

تجدر الإشارة إلى أن جميع مقاييس التشتت هي قيم موجبة، وذلك شرط أساسي يجب توفره في جميع مقاييس التشتت. لذا فان مقاييس التشتت لا تأخذ قيم سالبة أبدا بل تكون قيمها موجبة دوما أو مساوية للصفر فقط وذلك إذا كانت جميع قيم المتغير الكمي محل الدراسة متساوية، أي أنه لا يوجد تباين أو تشتت أصلا.

أولا: المدى : Range

المدى هو الفرق بين أكبر قيمة وأصغر قيمة في مجموعة ما، فمدى المجموعة الأولى 18- 9= 9 . **مثال (07)**:
بينما مدى المجموعة الثانية 19-2= 17 لاحظ أن المجموعة الثانية أكثر تشـتتاً من المجموعة الأولى، أما المدى لقيم معطاة في جدول توزيع تكراري فيحسـب من الفرق بين الحد الأعلى للفئة العليا والحد الأدنى للفئة الدنيا . ففي الجدول الخاص بالمثال رقم **(06)** المدى هو 34 = 20 - 54 .

يستخدم المدى الممثل للفرق بين أكبر قيمة وأصغر قيمة كمقياس بسيط وسطحي عن درجة تشتت قيم المتغير الكمي، ولكن لا يجب الأخذ بهذا المقياس والاعتماد عليه في العمليات الاستدلالية الإحصائية حيث أنه يتأثر بشدة بالقيم المتطرفة بالإضافة إلى عدم استخدامه لباقي قيم المتغير الكمي، وكمقياس أدق يعمل في حال وجود قيم متطرفة، يمكن استخدام الانحراف الربيعي والذي يعتمد على ترتيب القيم كما هو معمول به في عملية حساب الوسيط، في المقابل عندما لا تكون مشكلة القيم المتطرفة حاضرة فانه يمكن استخدام الانحراف المتوسط أو الانحراف المعياري كمقياس للتشتت، حيث يتم استخدام كافة القيم في عملية حساب المقاييس السابقة.

ثانيا: الإنحراف المعياري والتباين Standard Deviation and variance

يعتبر الانحراف المعياري والتباين من أهم مقاييس التشتت الإحصائية، ويرتبط المقياسين بعلاقة رياضية قوية، حيث يمكن دوما الحصول على المقياس الآخر في حال معرفة قيمة أحدهما، يرمز للتباين بالرمز σ^2 في حال الحصول على قيمته من خلال تغطية مجتمع الدراسة، بينما يتم استخدام الرمز S^2 للدلالة على مقدر التباين المحصل من خلال بيانات عينة عشوائية مسحوبة من مجتمع الدراسة، وبأخذ الجذر التربيعي للتباين يتم الحصول على قيمة الانحراف المعياري وذلك في الحالتين، حالة المجتمع وحالة العينة،

$$\sigma = \sqrt{\sigma^2} \qquad \text{or} \qquad S = \sqrt{S^2}$$

وبحكـم العلاقـة الرياضـية القويـة بـين كـل مـن التبـاين والانحـراف المعيـاري فإنـه يمكـن اعتبارهمـا وجهين لعملة واحدة لهما نفس الأهمية.

يعتمـد الانحـراف المعيـاري والتبـاين علـى فكـرة تربيـع الفـروق بـين قـيم المتغـير الكمـي X ووسـطها الحسـابي، وتتأسـس فكـرة تربيـع فـروق القـيم عـن متوسـطها الحسـابي مـن مشـكلة تتعلـق بمجمـوع تلـك الفـروق، فقـد لاحظنـا في الجـزء المتعلـق بالمتوسـط الحسـابي أن مجمـوع الفـروق (الانحرافـات) دائمـا وأبـدا تسـاوي 0، وبمـا التعريـف الـدقيق للانحـراف المعيـاري هـو في الحقيقـة متوسـط الانحرافـات، ولأن مجمـوع الإنحرافـات أصـلا يسـاوي صـفر فقـد تم تربيـع هـذه الانحرافـات للـتخلص مـن الإشـارة السـالبة للفـروق الـتي تقـل عـن قيمـة المتوسـط الحسـابي، لـذلك فـإن حسـاب متوسـط مربـع الفـروق يعطينـا قيمـة التبـاين بينمـا جذر القيمة المتحصل عليها يعطينا الإنحراف المعياري.

مثال (08): انطلاقا من القيم التالية 9، 10، 10، 11، 12، 12، 14، 18

1. حساب المتوسط الحسابي: $\bar{x} = \frac{1}{8}[9+10+10+11+12+14+18] = \frac{96}{8} = 12$
2. حساب مجموع فروق القيم عن المتوسط (الانحرافات) $\sum (x_1-\bar{x})+(x_2-\bar{x})+(x_3-\bar{x})...$

نجد: (12-18) + (12-14) + (12-12) + (12-11) + (12-10) + (12-10) (12-9)

= (+6) + (+2) + (0) + (-1) + (-2) +(-2) +(-3) =0

وهو ما يعني أن مجموع الإنحرافات يســـاروي 0 لذلك نلجأ إلى تربيع تلك الانحرافات وهو ما يتضـــمنه قانون حساب قيمة الانحراف المعياري والتباين بما أن:

الانحراف المعياري = $\sqrt{\text{التباين}}$

1. تحسب قيمة التباين للبيانات غير المبوبة من القانون التالي:

$$S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1} \qquad [10]$$

ويستخدم الباحثون النسخة المختصرة للقانون

$$S^2 = \frac{n\sum X_i^2 - (\sum X_i)^2}{n(n-1)}. \qquad [11]$$

مثال (09): انطلاقا من القيم التالية: 18، 10، 11، 12، 6، 8، 5 أوجد قيمة التباين باستخدام المعادلتين [10] و [11]

الحل وفق المعادلة[10]

- حساب قيمة المتوسط الحسابي:

$$\bar{x}=\frac{1}{7}\sum_{i=1}^{7}x_i \quad \bar{x}=\frac{1}{7}[5+8+6+12+11+10+18]=\frac{70}{7}=10$$

○ حساب مربعات انحرافات القيم عن متوسطها الحسابي

X	5	8	6	12	11	10	18	$\sum = 70$
$(X_i - \bar{X})$	5-	2-	4-	2+	1+	0	8+	$\sum = 0$
$(X_i - \bar{X})^2$	25	4	16	4	1	0	64	$\sum = 114$

$$S^2=\frac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1} = \frac{114}{7-1}=19$$

بما ان قيمة التباين هي 19

فإن قيمة الإنحراف المعياري هي $S=\sqrt{19}=4.35$

الحل باستخدام المعادلة [11]

X	5	8	6	12	11	10	18	$\Sigma X = 70$	$(\Sigma X)^2 = 4900$
X²	25	64	36	144	121	100	324	$\Sigma X^2 = 814$	

$$S^2=\frac{n\Sigma X_i^2-(\Sigma X_i)^2}{n(n-1)} = \frac{7(814)-(70)^2}{7(7-1)}=\frac{5698-4900}{42}=19$$

بما ان قيمة التباين هي 19فإن قيمة الإنحراف المعياري هي $S=\sqrt{19}=4.35$

2. حساب قيمة التباين والانحراف المعياري للبيانات المبوبة:

$$\boxed{S^2=\frac{\sum_{i=1}^{k}(X_i-\bar{X})^2 f_i}{n-1}} \qquad \textbf{[12]}$$

$$\boxed{S^2=\frac{n\Sigma f_i X_i^2-(\Sigma f_i X_i)^2}{n(n-1)}} \qquad \textbf{[13]}$$

مثال (010): انطلاقا من القيم الواردة في الجدول أدناه أوجد قيمة التباين والانحراف المعياري باستخدام المعادلة [13]

المجموع	34 – 30	29 – 25	24 – 20	19 – 15	14 – 10	9 – 5	4 – 0	الفئات
/	32	27	22	17	12	7	2	مراكز الفئات X
56	5	12	7	5	9	10	8	التكرار fi
917	160	324	154	85	108	70	16	X fi

	1024	729	484	289	144	49	4	X^2
17519	5120	8748	3388	1445	1296	490	32	$X^2 fi$

$$S^2 = \frac{n\sum f_i X_i^2 - (\sum f_i X_i)^2}{n(n-1)} \qquad \frac{56(17519)-(917)^2}{56(56-1)} = \frac{981064-840889}{3080} = 45.51$$

قيمة التباين هي 45.51 وعليه فإن قيمة الانحراف المعياري $S = \sqrt{45.51} = 6.74$

ملاحظات هامة :

(1)–إن السبب في القسمة على (n-1) عوضاً عن n لأن هناك (n-1) إنحرافاً مستقلاً من الشكل $x_i - \bar{x}$. ولأن مجموع هذه الانحرافات يساوي الصفر دوماً فإن أيا منها يساوي مجموع كل البقية بإشارة سالبة و أي منها يعطى بدلالة مجموع القيم الأخرى وبإشارة معاكسة . ولتوضيح هذه الفكرة تصور أن لدينا ثلاث بيانات x_1 , x_2 , x_3 ولدينا $\bar{x} = \frac{1}{3}(x_1 + x_2 + x_3)$.

نعلم أن $(x_1 - \bar{x}) + (x_2 - \bar{x}) + (x_3 - \bar{x}) = 0$

وبالتالي يمكن التعبير عن أي منهم وليكن الأول ب $x_1 - \bar{x} = -[(x_2 - \bar{x}) + (x_3 - \bar{x})]$

أو $x_2 - \bar{x} = -[(x_1 - \bar{x}) + (x_3 - \bar{x})]$

أو $x_3 - \bar{x} = -[(x_1 - \bar{x}) + (x_2 - \bar{x})]$

أي يمكن تمثيل أي انحراف بدلالة الاثنين الآخرين وحيث لدينا n = 3 فسنقسم على 2=1-3 .