المحور الأول: مفاهيم الأساسية الإحصاء والاحتمالات

قبل الخوض في طرق الاقتصاد القياس لبد من التذكير من بعض المفاهيم الأساسية الخاصة بالإحصاء و الاحتمالات وسوف نركز على الوسط الحسابي ، التباين والانحراف المعياري، وبعض التوزيعات الاحتمالية والتي لها علاقة وطيدة بمقياس الاقتصاد القياسي ، لان الطلبة سبق لهم وان تطرقوا لهده المفاهيم في مقياس الإحصاء الوصفي والإحصاء الرياضي

1 - الوسط الحسابي

من أهم مقاييس النزعة المركزية ، وأكثرها استخداما في النواحي التطبيقية ، ويمكن حسابه للبيانات المبوبة وغير المبوبة ، كما يلي :

1-1- الوسط الحسابي للبيانات غير المبوبة

يعرف الوسط الحسابي بشكل عام على أنه مجموع القيم مقسوما على عددها . فإذا كان لدينا n من القيم ، ويرمز لها بالرمز : $x_1, x_2, ..., x_n$.

فإن الوسط الحسابي لهذه القيم ، ونرمز له بالرمز $\overline{x}$ يحسب بالمعادلة التالية :

$$\text{الوسط الحسابي} = \frac{\text{مجموع القيم}}{\text{عدد القيم}}$$

$$\overline{x} = \frac{x_1 + x_2 + ... + x_n}{n} = \frac{\sum_{i=1}^{n} x_i}{n}$$

حيث يدل الرمز Σ على المجموع .

مثـــال 01

فيما يلي درجات 8 طلاب في مقرر 122 إحصاء تطبيقي .

34 32 42 37 35 40 36 40

والمطلوب إيجاد الوسط الحسابي لدرجة الطالب في الامتحان .

الحـــل

لإيجاد الوسط الحسابي للدرجات تطبق المعادلةالسابقة كما يلي:

$$\overline{x} = \frac{x_1 + x_2 + ... + x_n}{n}$$

$$= \frac{34+32+42+37+35+40+36+40}{8} = \frac{296}{8} = 37$$

أي أن الوسط الحسابي لدرجة الطالب في اختبار مقرر 122 إحصاء يساوي 37 درجة

1-2- الوسط الحسابي للبيانات المبوبة

من المعلوم أن القيم الأصلية ، لا يمكن معرفتها من جدول التوزيع التكراري ، حيث أن هذه القيم موضوعة في شكل فئات ، ولذا يتم التعبير عن كل قيمة من القيم التي تقع داخل حدود الفئة

بمركز هذه الفئة ، ومن ثم يؤخذ في الاعتبار أن مركز الفئة هو القيمة التقديرية لكل مفردة تقع في هذه الفئة.

فإذا كانت k هي عدد الفئات ، وكانت $x_1, x_2, ..., x_k$ هي مراكز هذه الفئات، $f_1, f_2, ..., f_k$ هي التكرارات ، فإن الوسط الحسابي يحسب بالمعادلة التالية:

$$\bar{x} = \frac{x_1 f_1 + x_2 f_2 + ... + x_k f_k}{f_1 + f_2 + ... + f_k} = \frac{\sum_{i-1}^{k} x_i f_i}{\sum_{i-1}^{k} f_i}$$

مثـــال 02:

الجدول التالي يعرض توزيع 40 تلميذ حسب أوزانهم .

فئات الوزن	32-34	34-36	36-38	38-40	40-42	42-44
عدد التلاميذ	4	7	13	10	5	1

والمطلوب إيجاد الوسط الحسابي.

الحـــــل

لحساب الوسط الحسابي باستخدام المعادلة رقم (3-2) يتم إتباع الخطوات التالية :

1- إيجاد مجموع التكرارات $\sum f$. 2- حساب مراكز الفئات x .

3- ضرب مركز الفئة في التكرار المناظر له (xf)، وحساب المجموع $\sum xf$

4- حساب الوسط الحسابي بتطبيق المعادلة رقم (3-2) .

فئات الوزن (C)	التكرارات f	مراكز الفئات x	$x\ f$
32-34	4	(32+34)÷2=33	4×33=132
34-36	7	35	7×35=245
36-38	13	37	13×37=481
38-40	10	39	10×39=390
40-42	5	41	5×41=205
42-44	1	43	1×43=43
المجموع	40		1496

إذا الوسط الحسابي لوزن التلميذ هو : $\bar{x} = \frac{\sum_{i=1}^{6} x_i f_i}{\sum_{i=1}^{6} f_i} = \frac{1496}{40} = 37.4\ k.g$

أي أن متوسط وزن التلميذ يساوي 37.4 k.g

1-3- خصائص الوسط الحسابي: يتصف الوسط الحسابي بعدد من الخصائص ، ومن هذه الخصائص ما يلي :

❖ الوسط الحسابي للمقدار الثابت يساوى الثابت نفسه ، أي أنه إذا كانت قيم x هي : $x : a, a, ..., a$ ، فإن الوسط الحسابي هو:

$$\overline{x} = \frac{a + a + \ldots + a}{n} = \frac{na}{n} = a$$

ومثال على ذلك ، لو اخترنا مجموعة من 5 طلاب ، ووجدنا أن كل طالب وزنه 63 كيلوجرام ، فإن متوسط وزن الطالب في هذه المجموعة هو :

$$\overline{x} = \frac{63+63+63+63+63}{5} = \frac{315}{5} = 63\ k.g$$

❖ مجموع انحرافات القيم عن وسطها الحسابي يساوى صفرا ، ويعبر عن هذه الخاصية بالمعادلة .

$$\sum(x-\overline{x}) = 0$$

ويمكن التحقق من هذه الخاصية باستخدام بيانات مثال 01 ، نجد أن درجات الطلاب هي : 34, 32, 42, 37, 35, 40, 36, 40 ، والوسط الحسابي للدرجة هو $\overline{x} = 37$ ، إذا :

x	34	32	42	37	35	40	36	40	296
$(x-\overline{x})$	34-37	32-37	42-37	37-37	35-37	40-37	36-37	40-37	0
$(x-37)$	-3	-5	5	0	-2	3	-1	3	

أي أن : $\sum(x-37) = 0$

- إذا أضيف مقدار ثابت إلى كل قيمة من القيم ، فإن الوسط الحسابي للقيم المعدلة (بعد الإضافة) يساوى الوسط الحسابي للقيم الأصلية (قبل الإضافة) مضافا إليها هذا المقدار الثابت . فإذا كانت القيم هي : $x_1, x_2, ..., x_n$ ، وتم إضافة مقدار ثابت (a) إلى كل قيمة من القيم ، ونرمز للقيم الجديدة بالرمز y ، أي أن $y = x + a$ ، فإن : الوسط الحسابي لقيم y (القيم بعد الإضافة) هو:

$$\overline{y} = \overline{x} + a$$

حيث أن $\overline{y}$ هو الوسط الحسابي للقيم الجديدة ، ويمكن التحقق من هذه الخاصية باستخدام بيانات مثال رقم 01 .

إذا قرر المصحح إضافة 5 درجات لكل طالب ، فإن الوسط الحسابي للدرجات المعدلة يصبح قيمته {(37+5)=42} ، والجدول التالي يبين ذلك .

x	34	32	42	37	35	40	36	40	296
$y=(x+5)$	34+5	32+5	42+5	37+5	35+5	40+5	36+5	40+5	336
	39	37	47	42	40	45	41	45	

نجد أن مجموع القيم الجديدة هو : $\sum y = 336$ ، ومن ثم يكون الوسط الحسابي للقيم الجديدة هو :

$$\bar{y} = \frac{\sum y}{n} = \frac{336}{8} = 42 \quad \rightarrow (\bar{x} + 5 = 37 + 5 = 42)$$

- إذا ضرب مقدار ثابت (a) في كل قيمة من القيم ، فإن الوسط الحسابي للقيم المعدلة (القيم الناتجة بعد الضرب) يساوى الوسط الحسابي للقيم الأصلية (القيم بعد التعديل) مضروبا في هذا المقدار الثابت . أى أنه إذا كان : $y = a\ x$ ، ويكون الوسط الحسابي للقيم الجديدة y هو :

$$\bar{y} = a\ \bar{x}$$

ويمكن للطالب أن يتحقق من هذه الخاصية باستخدام نفس بيانات المثال السابق . فإذا كان تصحيح الدرجة من 50 ، وقرر المصحح أن يجعل التصحيح من 100 درجة ، بمعنى أنه سوف يضرب كل درجة في قيمة ثابتة (a=2) ، ويصبح الوسط الحسابي الجديد هو : $\bar{y} = a\ \bar{x} = 2(37) = 74$

❖ مجموع مربعات انحرافات القيم عن وسطها الحسابي أقل ما يمكن ، أي أن:

$$\sum (x-\bar{x})^2 < \sum (x-a)^2 \quad if \quad a \neq \bar{x}$$

وفي المثال السابق فإن : $\sum (x-37)^2 < \sum (x-a)^2$ لجميع قيم $a \neq 37$

1-4- الوسط الحسابي المرجح: في بعض الأحيان يكون لكل قيمة من قيم المتغير أهمية نسبية تسمى أوزن ، أو ترجيحات ، وعدم أخذ هذه الأوزان في الاعتبار عند حساب الوسط الحسابي ، تكون القيمة المعبرة عن الوسط الحسابي غير دقيقة ، فمثلا لو أخذنا خمسة طلاب ، وسجلنا درجات هؤلاء الطلاب في مقرر الإحصاء التطبيقي ، وعدد ساعات الاستذكار في الأسبوع .

مسلسل	1	2	3	4	5	مج
x (الدرجة)	23	40	36	28	46	173
w (عدد ساعات الاستذكار)	1	3	3	2	4	13

نجد أن الوسط الحسابي غير المرجح للدرجة الحاصل عليها الطالب هي :

$$\bar{x} = \frac{\sum x}{n} = \frac{23+40+36+28+46}{5} = \frac{173}{5} = 346$$

وإذا أردنا أن نحسب الوسط الحسابي للدرجات x المرجحة بعدد ساعات الاستذكار w ، يتم تطبيق المعادلة التالية :

$$(\overline{w})=\frac{\sum xw}{\sum w}=\frac{23\times1+40\times3+36\times3+28\times2+46\times4}{1+3+3+2+4}$$
$$=\frac{23+120+108+56+184}{13}=\frac{491}{13}=37.769$$

وهذا الوسط المرجح أكثر دقة من الوسط الحسابي غير المرجح .

إذا الوسط الحسابي المرجح $(\overline{w})$ يحسب بتطبيق المعادلة التالية :

$$\boxed{(\overline{w})=\frac{\sum xw}{\sum w}}$$

2- مزايا وعيوب الوسط الحسابي: يتميز الوسط الحسابي بالمزايا التالية :

- أنه سهل الحساب .
- يأخذ في الاعتبار كل القيم .
- أنه أكثر المقاييس استخداما وفهما .

ومن عيوبه .

- أنه يتأثر بالقيم الشاذة والمتطرفة .
- يصعب حسابه في حالة البيانات الوصفية .
- يصعب حسابه في حالة الجداول التكرارية المفتوحة .

1-5- التوقع الرياضي للمتغيرات العشوائية: يعرف التوقع الرياضي لمتغيرة عشوائية متقطعة كما يلي:

$$\boxed{E(X)=\sum_{i=1}^{n}x_iP(X=x_i)=\sum_{i=1}^{n}x_i\,f(x_i)}$$

و يعرف التوقع الرياضـــي لمتغيرة عشوائية مستمرة كما يلي:

$$E(X)=\int_{-\infty}^{+\infty}xf(x)dx$$

نرمز للتوقع أحيانا ب μ أو μ_x.

مثال01: نلقي قطعة نقدية 4 مرات. أحسب العدد المتوقع من المرات التي نحصل فيها على وجه.

Xعدد مرات وجه	0	1	2	3	4	المجموع
P(X)	$(1/2)^4$	$4\ (1/2)^4$	$6\ (1/2)^4$	$4\ (1/2)^4$	$(1/2)^4$	16/16 = 1
XP(X)	0	$4\ (1/2)^4$	$12\ (1/2)^4$	$12\ (1/2)^4$	$4\ (1/2)^4$	32/16 = 2

$$E(X)=\sum_{i=1}^{n}x_iP(X=x_i)=32/16=2$$

العدد المتوقع هو مرتين من بين 4 رميات

مثال 02: أوجد التوقع الرياضي للمتغيرة ذات دالة الكثافة التالية:

$$f(x)=\begin{cases}\frac{1}{2}x & 0<x<2\\ 0 & \text{sinon}\end{cases}$$

$$E(x)=\int_{-\infty}^{\infty}xf(x)dx=\int_{-\infty}^{0}x.0.dx+\int_{0}^{2}x(\frac{1}{2}x)dx+\int_{2}^{\infty}x.0.dx$$

$$E(x) = 0 + \frac{1}{2}\left[\frac{x^3}{3}\right]_0^2 + 0 = \frac{4}{3}$$

2 - :التباين

هو أحد مقاييس التشتت ، وأكثرها استخداما في النواحي التطبيقية ، ويعبر عن متوسط مربعات انحرافات القيم عن وسطها الحسابي.

أولا: التباين في المجتمع (σ^2)

إذا توافر لدينا قراءات عن كل مفردات المجتمع ، ولتكن: $x_1, x_2, \ldots, x_N$ ، فإن التباين في المجتمع ، ويرمز له بالرمز σ^2 يحسب باستخدام المعادلة التالية :

$$\sigma^2 = \frac{\sum (x-u)^2}{N}$$

حيث أن μ هو الوسط الحسابي في المجتمع ، أى أن : $\mu = \sum x/N$.

مثـــال:

مصنع لتعبئة المواد الغذائية ، يعمل به 15 عامل ، وكانت عدد سنوات الخبرة لهؤلاء العمال كما يلي :

5 13 7 14 12 9 6 8 10 13 14 6 11 12 10

بفرض أن هذه البيانات تم جمعها عن كل مفردات المجتمع ، فأوجد التباين لعدد سنوات الخبرة .

الحـــل

لحساب تباين سنوات الخبرة في المجتمع ، يتم استخدام المعادلة (4-6).

- الوسط الحسابي في المجتمع μ

$$\mu = \frac{1}{N}\sum x$$

$$= \frac{1}{15}(5+13++7+\ldots+12+10) = \frac{1}{15}(150) = 10$$

- حساب مربعات الانحرافات $\sum (x-\mu)^2$

بما أن: $\sum (x-\mu)^2 = 130$

إذا تباين سنوات الخبرة للعمال في المصنع هو :

$$\sigma^2 = \frac{\sum (x-u)^2}{N} = \frac{130}{15} = 8.67$$

سنوات الخبرة x	$(x-\mu)$	$(x-\mu)^2$
5	5-10 = -5	25
13	3	9
7	-3	9
14	4	16
12	2	4
9	-1	1
6	-4	16
8	-2	4
10	0	0
13	3	9
14	4	16
6	-4	16
11	1	1
12	2	4
10	0	0
150	0	130

ويمكن تبسيط المعادلة (4-6) في صورة أخرى كما يلي :

يمكن فك المجموع $\sum (x-\mu)^2$ كالتالي :

$$\begin{aligned}\sum (x-\mu)^2 &= \sum \left(x^2 - 2x\mu + \mu^2 \right) \\ &= \sum x^2 - 2\mu \sum x + \sum \mu^2 \\ &= \sum x^2 - 2N\mu^2 + N\mu^2 \\ &= \sum x^2 - N\mu^2\end{aligned}$$

ومن ثم يكتب تباين المجتمع على الصورة التالية :

$$\sigma^2 = \frac{\sum x^2 - N\mu^2}{N} = \frac{1}{N}\sum x^2 - \mu^2$$

إذا التباين في المجتمع يمكن صياغته كالتالي .

$$\boxed{\sigma^2 = \frac{1}{N}\sum x^2 - \mu^2}$$

وبالتطبيق على المثال ، نجد أن أننا نحتاج إلى المجموعين : $\sum x^2$, $\sum x$ ، ويتم عمل الآتي :

سنوات الخبرة x	x^2
5	25
13	169
7	49
14	196
12	144
9	81
6	36
8	64
10	100
13	169
14	196
6	36
11	121
12	144
10	100
150	1630

$$\sum x = 150 \ , \ \sum x^2 = 1630$$

$$\mu = \frac{1}{N}\sum x = \frac{1}{15}(150) = 10$$

إذا التباين هو

$$\sigma^2 = \frac{1}{N}\sum x^2 - \mu^2$$

$$= \frac{1}{15}1630 - 10^2 = 108.67 - 100 = 8.67$$

وهي نفس النتيجة التي تم الحصول عليها باستخدام الصيغة السابق.

2-1- التباين في العينة (s^2)

في كثير من الحالات يكون تباين المجتمع σ^2 غير معلوم، وعندئذ يتم سحب عينة من هـذا المجتمع ، ويحسب التباين من بيانات العينة كتقدير لتباين المجتمع ، فإذا كانت قراءات عينة عشوائية حجمها n هي ، $x_1, x_2, \ldots, x_n$ ، فإن تباين العينة ويرمز له بالرمز s^2 هو:

$$s^2 = \frac{\sum (x - \bar{x})^2}{n-1}$$

حيث أن $\bar{x}$ هو الوسط الحسابي لقراءات العينة ، أي أن : $\bar{x} = \sum x / n$ ، وتبـاين العينـة المبين بالمعادلة هو التقدير غير المتحيز لتباين المجتمع .

مثـال

في المثال السابق ، إذا تم سحب عينة من عمال المصنع حجمها 5 عمـال ، وسـجل عـدد سنوات الخبرة ، وكانت كالتالي .

8 13 10 5 9

احسب تباين سنوات الخبرة في العينة .

الحـــــل: لحساب التباين في العينة يتم تطبيق المعادلة (4-8)، ويتبع الآتي :

- الوسط الحسابي في العينة :

$$\bar{x} = \frac{1}{n}\sum x = \frac{1}{5}(8+13+10+5+9) = \frac{1}{5}(45) = 9$$

- حساب مربعات الانحرافات $\sum (x - \bar{x})^2$

سنوات الخبرة x	8	13	10	5	9	45
$(x - \bar{x})$	-1	4	1	-4	0	0
$(x - \bar{x})^2$	1	16	1	16	0	34

أي أن : $\sum (x - \bar{x})^2 = 34$ ،

- إذا تباين سنوات الخبرة في العينة قيمته هي :

$$s^2 = \frac{\sum (x - \bar{x})^2}{n - 1} = \frac{34}{(5 - 1)} = \frac{34}{4} = 8.5$$

- في هذه الحالة يمكن القول بأن تباين العينة 8.5، وهو في نفس الوقت تقدير غير متحيز لتبـــاين المجتمع .

تبسيط العمليات الحسابية

يمكن تبسيط الصيغة الرياضية لتباين العينة الموضحة بالمعادلة (4-8) إلى صيغة سهلة يمكن التعامل معها، وخاصة إذا كانت البيانات تحتوي على قيم كسرية، ولاستنتاج هذه الصيغة يتم إتبـــاع الآتي.

يمكن فك المجموع $\sum (x-\bar{x})^2$ كالتالي:

$$\begin{aligned}\sum (x-\bar{x})^2 &= \sum \left(x^2 - 2x\bar{x} + \bar{x}^2 \right) \\ &= \sum x^2 - 2\bar{x}\sum x + \sum \bar{x}^2 \\ &= \sum x^2 - 2n\bar{x}^2 + n\bar{x}^2 \\ &= \sum x^2 - n\bar{x}^2\end{aligned}$$

ويكتب تباين العينة على الصورة التالية :

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - n\bar{x}^2 \right)$$

إذا التباين في العينة يمكن صياغته كالتالي .

$$\boxed{s^2 = \frac{1}{n-1}\left(\sum x^2 - n\bar{x}^2 \right)}$$

كما يمكن إثبات أن المعادلة (4-9) تأخذ الشكل التالي:

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)$$

وبالتطبيق على بيانات المثال السابق ، نجد أن :

سنوات الخبر x	8	13	10	5	9	45
x^2	64	169	100	25	81	439

- تباين العينة باستخدام المعادلة (4-9) هو :

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - n\bar{x}^2\right)$$
$$= \frac{1}{5-1}\left(439 - 5(9)^2\right) = \frac{1}{4}(34) = 8.5$$

- وباستخدام المعادلة (4-10) نجد أن:

$$s^2 = \frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)$$
$$= \frac{1}{5-1}\left(439 - \frac{(45)^2}{5}\right) = \frac{1}{4}(439 - 405) = \frac{1}{4}(34) = 8.5$$

2-2- التباين للمتغيرات العشوائية: يعرف التباين لمتغيرة عشوائية كما يلي:

$$\sigma_X^2 = Var(X) = E\left[(X - \mu)^2\right]$$

و الانحراف المعياري هو جذر التباين. $\sigma = \sqrt{V(X)} = \sqrt{\sigma^2}$.

في حالة المتغيرة العشوائية المتقطعة :

$$V(X) = \sum_x (x-\mu)^2 f(x)$$

في حالة المتغيرة العشوائية مستمرة :

$$V(X) = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx$$

مثال. نلقي قطعة نقدية مرتين. نسمي X عدد مرات الحصول على صورة، أحسب V(X) .

X	0	1	2	المجموع
P(X = x)	1/4	1/2	1/4	1
X*P(X)	0	1/2	1/2	E(X) = μ =1
$(X-\mu)^2$	1	0	1	
$(X-\mu)^2$ * P(X)	1/4	0	1/4	V(X) = 1/2

مثال 2. لتكن X م ع ذات دالة الكثافة التالية؛ أحسب تباين X.

$$f(x) = \begin{cases} \frac{1}{2}x & 0 < x < 2 \\ 0 & \text{sinon} \end{cases}$$

$\mu = 4/3$ $\qquad \sigma^2 = \int_{-\infty}^{+\infty} (x-\mu)^2 f(x)dx.$

$$\sigma^2 = \int_{-\infty}^{0}\left(x-\frac{4}{3}\right)^2 * 0dx + \int_{0}^{2}\left(x-\frac{4}{3}\right)^2 \frac{x}{2}dx + \int_{2}^{+\infty}\left(x-\frac{4}{3}\right)^2 0dx = 0 + \frac{1}{2}x\left(x^2 - \frac{8x}{3} + \frac{16}{9}\right)_0^2 + 0$$

$$\sigma^2 = \frac{1}{2}\left(x^3 - \frac{8x^2}{3} + \frac{16x}{9}\right)_0^2 = \frac{1}{2}\left(8 - \frac{32}{3} + \frac{32}{9}\right) = \frac{4}{9}$$

3- الانحراف المعياري

عند استخدام التباين كمقياس من مقاييس التشتت، نجد أنه يعتمد علي مجموع مربعات الانحرافات، ومن ثم لا يتمشى هذا المقياس مع وحدات قياس المتغير محل الدراسة ، ففي المثال السابق ، نجد أن تباين سنوات الخبرة في العينة 8.5، فليس من المنطق عند تفسير هذه النتيجة أن نقول ، " تباين سنوات الخبرة هو 8.5 سنة تربيع "، لأن وحدات قياس المتغير هو عدد السنوات، من أجل ذلك لجأ الإحصائيين إلى مقياس منطقي يأخذ في الاعتبار الجذر التربيعي للتباين ، لكي يناسب وحدات قياس المتغير، وهذا المقياس هو الانحراف المعياري.

إذا الانحراف المعياري ، هو الجذر التربيعي الموجب للتباين ، أي أن:

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

ومثال على ذلك :

- في مثال (4-7) نجد أن الانحراف المعياري لسنوات الخبرة لعمال المصنع (المجتمع) ، ويرمز له بالرمز (σ) هو :

$$\sigma = \sqrt{\frac{1}{N}\sum x^2 - \mu^2}$$

$$= \sqrt{\frac{1}{15}1630 - 10^2} = \sqrt{8.67} = 2.94$$

في هذه الحالة ، يكون الانحراف المعياري لسنوات الخبرة في المجتمع هو 2.94 سنة .

- في مثال السابق نجد أن الانحراف المعياري لسنوات الخبرة لعمال العينة ، ويرمز له بالرمز s ، هو :

$$s = \sqrt{\frac{1}{n-1}\left(\sum x^2 - \frac{(\sum x)^2}{n}\right)}$$

$$= \sqrt{\frac{1}{5-1}\left(439 - \frac{(45)^2}{5}\right)} = \sqrt{\frac{1}{4}(439-405)} = \sqrt{\frac{1}{4}(34)} = 2.92$$

أي أن الانحراف المعياري لسنوات الخبرة في العينة هو 2.92 سنة .

3-1- الانحراف المعياري في حالة البيانات المبوبة

إذا كانت بيانات الظاهرة ، مبوبة في جدول توزيع تكراري ، فإن الانحراف المعياري يحسب

بتطبيق المعادلة التالية .

$$s = \sqrt{\frac{\sum (x - \bar{x})^2 f}{n-1}}$$

or

$$s = \sqrt{\frac{\sum x^2 f - \frac{(\sum xf)^2}{n}}{n-1}}$$

حيث أن f هو تكرار الفئة ، x هو مركز الفئة ، $\overline{x}$ هو الوسط الحسابي $(\sum xf/n)$ ، n هي مجموع التكرارات $(n = \sum f)$ ، والمقدار الذي تحت الجذر يعبر عن التباين (s^2) .

مثـــال: يبين الجدول التكراري التالي توزيع **40** أسرة حسب الإنفاق الشهري بالألف الدينار.

الإنفاق	**2 - 5**	**5 - 8**	**8 - 11**	**11 – 14**	**14 – 17**
عدد الأسرة	**1**	**8**	**13**	**10**	**8**

احسب الانحراف المعياري للإنفاق الشهري للأسرة ، ثم قارن بين الانحراف المتوسط ، والانحراف المعياري للإنفاق الشهري للأسرة

الحـــــــــل:لحساب الانحراف المعياري للإنفاق الشهري ، تســتخدم المعادلــة رقـــم (4-12)، وسوف نطبق الصيغة الثانية ، ولذا نكون جدول لحساب المجموعين : $\sum x^2 f$, $\sum xf$.

الإنفاق	**عدد الأسر** f	**مركز الفئة** x	xf	$x^2 f$
2-5	1	3.5	3.5	12.25
5-8	8	6.5	52	338
8-11	13	9.5	123.5	1173.25
11-14	10	12.5	125	1562.5
14-17	8	15.5	124	1922
sum	40		428	5008

$$n = \sum f = 40$$
$$\sum xf = 428$$
$$\sum x^2 f = 5008$$

وبتطبيق المعادلة ، نجد أن الانحراف المعياري قيمته هي :

$$s = \sqrt{\frac{\sum x^2 f - \frac{(\sum xf)^2}{n}}{n-1}}$$

$$= \sqrt{\frac{5008 - \frac{(428)^2}{40}}{40-1}} = \sqrt{\frac{5008 - 4579.6}{39}}$$

$$= \sqrt{10.984615} = 3.314$$

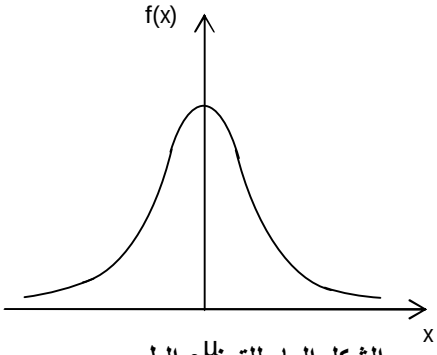
أي أن الانحراف المعياري للإنفاق الشهري **3.314** دينار.

4- التوزيعات الاحتمالية

4-1- التوزيع الطبيعي: يعتبر التوزيع الطبيعي من أهم التوزيعات الاحتمالية في علم الإحصاء لأنه يمثل كثيراً من الظواهر التي تقابلنا في الحياة العملية مثل الأطوال و الأوزان و الأعمار ودرجات الحرارة والدخول الشهرية ... و غيرها من الظواهر المتصلة . ولو مثلنا هذه البيانات في معلم متعامد متجانس لكان المنحنى الذي يمثل النسبة، أو ما يمكن أن نسميه الاحتمال، ذا شكل جرسي متماثل حول المتوسط وهي صفات التوزيع الطبيعي ويمكن توضيح ذلك بالشكل العام لتوزيع الطبيعي

تكتب دالة الكثافة لمنحنى للتوزيع الطبيعي كما يلي:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(1/2)\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$



الشكل العام للتوزيع الطبيعي

حيث μ و σ هما على التوالي التوقع والانحراف المعياري. ونكتب $X \sim N(\mu, \sigma)$

دالة التوزيع (الدالة التجميعية) للتوزيع الطبيعي تكتب كما يلي:

$$F(x) = P(X \le x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{x} e^{-1/2\left(\frac{v-\mu}{\sigma}\right)^2} dv$$

المتغيرة المركزية أو المعيارية : تستخدم المتغيرة المعيارية $Z = (X-\mu)/\sigma$ لتكوين الجداول الإحصائية للاحتمالات:

$$F(z) = P(Z \le z) \text{ أو } P(0 \le Z \le z)،$$

حيث تسمح بكتابة الدالة f و F بدلالة مجهول واحد Z بدلا من 3 مجاهيل x و μ و σ وذلك كما يلي:

$$f(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \qquad -\infty < z < \infty$$

$$F(z) = P(Z \le z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{\frac{-u^2}{2}} du$$

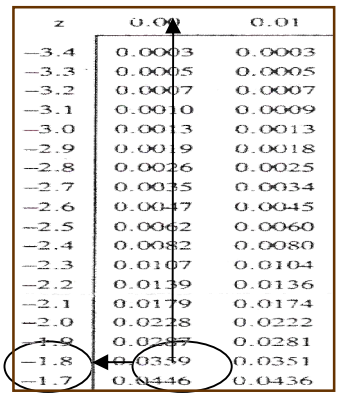
بالنظر إلى العلاقة الخطية بين المتغيرتين X و Z، فإن Z تتبع نفس توزيع X أي التوزيع الطبيعي. ونعلم أن:

V(Z) = 1 E(Z) = 0

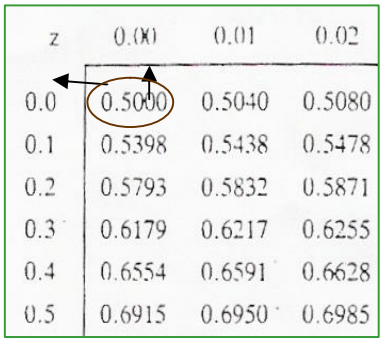
مثال: إذا كان Z متغير عشوائيا متصلا يتبع التوزيع الطبيعي القياسي فأوجد : $P(-1.8 \leq Z \leq 0)$

الحل:

$$P(-1.8 \leq Z \leq 0)$$
$$= P(Z \leq 0) - P(Z \leq -1.8)$$
$$= 0.5 - 0.359$$
$$= 0.4641$$



z	0.00	0.01
−3.4	0.0003	0.0003
−3.3	0.0005	0.0005
−3.2	0.0007	0.0007
−3.1	0.0010	0.0009
−3.0	0.0013	0.0013
−2.9	0.0019	0.0018
−2.8	0.0026	0.0025
−2.7	0.0035	0.0034
−2.6	0.0047	0.0045
−2.5	0.0062	0.0060
−2.4	0.0082	0.0080
−2.3	0.0107	0.0104
−2.2	0.0139	0.0136
−2.1	0.0179	0.0174
−2.0	0.0228	0.0222
−1.9	0.0287	0.0281
−1.8	0.0359	0.0351
−1.7	0.0446	0.0436



z	0.00	0.01	0.02
0.0	0.5000	0.5040	0.5080
0.1	0.5398	0.5438	0.5478
0.2	0.5793	0.5832	0.5871
0.3	0.6179	0.6217	0.6255
0.4	0.6554	0.6591	0.6628
0.5	0.6915	0.6950	0.6985

4-2- توزيع ك2 Distribution en Khi-carré (ou Khi-deux)**:** توزيع ك2 هو من أكثر التوزيعـــات استخداما في مجال اختبار الفروض بأنواعها، ويمكن تعريفه كما يلي:

لتكن $X_1, X_2, \ldots X_\nu$ ، متغيرات عشوائية مستقلة كل منها تتبع التوزيع الطبيعي المعياري ($\mu = 0, \sigma = 1$). المتغيرة

$$X = X_1^2 + X_2^2 + \ldots + X_\nu^2$$

لها دالة الكثافة التالية:

$$f(x) = \begin{cases} \dfrac{x^{(\nu/2)-1} e^{-x/2}}{2^{\nu/2}\Gamma(\nu/2)} & x > 0 \\ 0 & x \leq 0 \end{cases}$$

حيث $\Gamma(\alpha)$ هي الدالة قاما:

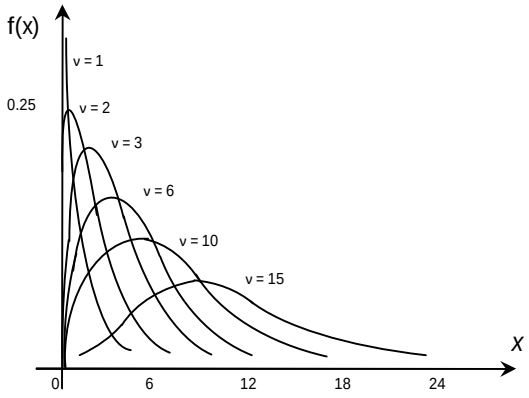
$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t} dt \qquad \alpha > 0$$

و نقول أن X تتبع التوزيع ك2 ب ν درجة حرية ونكتب

$$X \sim \chi^2_\nu .$$

الدالة التجميعية $F(\chi^2)$ تكتب كما يلي:

$$P(X^2 \leq x) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} \int_0^x u^{(\nu/2)-1} e^{-u/2} du & x \geq 0 \\ 0 & x < 0 \end{cases}$$



تدرج منحنى ك2 حسب درجة الحرية

3-4- توزيع ستيودنت

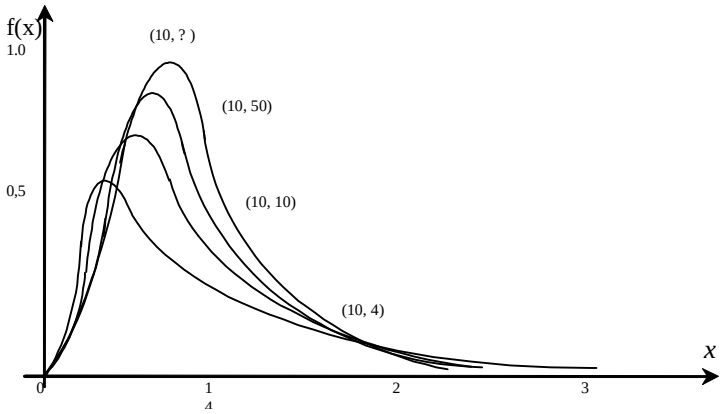
لتكن المتغيرتان العشوائيتان المستقلتان Y و Z حيث $Y \sim N(0,1)$ و $Z \sim \chi_\nu^2$ ؛ المتغيرة $T = \frac{Y}{\sqrt{Z/\nu}}$ تتبع توزيع لها دالة الكثافة التالية

$$f(t) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\nu\pi}\ \Gamma\left(\frac{\nu}{2}\right)} \left(1+\frac{t^2}{\nu}\right)^{-\left(\frac{\nu+1}{2}\right)} \qquad -\infty < t < \infty$$

$$\Gamma(\alpha) = \int_0^\infty t^{\alpha-1} e^{-t}\, dt \qquad \alpha > 0$$

و نقول أن المتغيرة X تتبع توزيع ستيودنت ب ν درجة حرية ونكتب: $T \sim t_\nu$

4-4- توزيع فيشر (F) Distribution F de Fisher-Snédecor



تدرج منحنى فيشر حسب درجة الحرية

ليكن لدينا المتغيرتان العشوائيتان المستقلتان $\chi_1 \sim \chi_{\nu_1}^2$ و $\chi_2 \sim \chi_{\nu_2}^2$. المتغيرة : $X = \frac{X_1/\nu_1}{X_2/\nu_2}$ لها دالة الكثافة:

$$f(x) = \begin{cases} \dfrac{\Gamma\left(\frac{\nu_1+\nu_2}{2}\right)}{\Gamma\left(\frac{\nu_1}{2}\right)\Gamma\left(\frac{\nu_2}{2}\right)}\ \nu_1^{\frac{\nu_1}{2}}\ \nu_2^{\frac{\nu_2}{2}}\ x^{(\frac{\nu_1}{2})-1}\ \left(\nu_2+\nu_1 x\right)^{-\frac{\nu_1+\nu_2}{2}} & x > 0 \\ 0 & x \le 0 \end{cases}$$

و نقول أن المتغيرة X تتبع توزيع فيشر ب ν_1 و ν_2 درجة حرية ونكتب: $X \sim F_{\nu1,\,\nu2}$