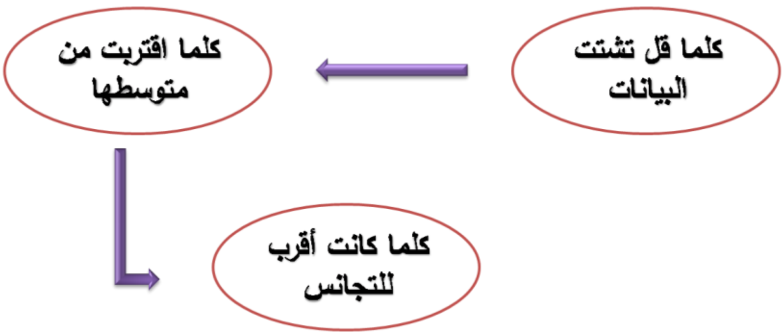
مقاييس التشتت

(المدى، الإنحراف المتوسط، التباين، الإنحراف المعياي)

تعريف التشتت

درجة التباعد أو التقارب التي تتجه بها البيانات الكمية للانتشار حول قيمة متوسطة (أحد مقاييس النزعة المركزية) تُسمى تشتت أو تغير البيانات. وتستخدم مقاييس التشتت في المقارنة بين مجموعات البيانات من حيث تشتتها.



هل يمكن الاكتفاء بالوسط الحسابي في وصف البيانات؟

إذا كان لدينا ٣ مجموعات من الطلاب، كل مجموعة مكونة من خمسة طلاب، وكانت درجاتهم في أحد المقررات كالتالي:

المجموعة الأولى	المجموعة الثانية	المجموعة الثالثة
5 , 5 , 5 , 5 , 5	3 , 4 , 5 , 6 , 7	1 , 2 , 5 , 8 , 9

وسطها الحسابي	وسطها الحسابي	وسطها الحسابي
5	**5**	**5**

المجموعات الثلاثة لها وسط حسابي 5 ، لكن في المجموعة الأولى: جميع القيم متساوية وتساوي الوسط 5 ، في حين تنتشر البيانات في المجموعة الثانية حول هذا الوسط بقدرٍ ما ، وفي المجموعة الثالثة تنتشر البيانات حول الوسط بقدرٍ آخر .

أي أن الوسط الحسابي وحده ليس كافياً وحده لوصف البيانات، ولكن لابد من وجود نوع آخر من المقاييس لرصد مدى تشتت البيانات عن تلك القيمة المتوسطة الممثلة للبيانات. هذا النوع من المقاييس هو ما نسميه بـ **مقاييس التشتت**

أولاً : المدى R :

البيانات الغير مبوبة

الفرق بين أكبر قيمة و أصغر قيمة في البيانات
و يرمز له بالرمز **R**

15 3 7 6 12 18 5 3 13 15 يكون المدى : $R = 18 - 3 = 15$

البيانات المبوبة

الفرق بين الحد الأعلى للفئة الأخيرة و الحد الأدنى للفئة الأولى

الفئة	العمر x
الأولى	$2 \le x < 6$
الثانية	$6 \le x < 12$
الثالثة	$12 \le x < 15$
الرابعة	$15 \le x < 18$

$R = 18 - 2 = 16$

الحد الأعلى للفئة الأخيرة — الحد الأدنى للفئة الأولى

وبالرغم من بساطة تحديده إلا أن له بعض العيوب:

- تأثره بالقيم المتطرفة

فمثلاً لمجموعة القيم : 15 3 7 6 12 18 5 3 13 15 يكون المدى $R = 18 - 3 = 15$

ولمجموعة القيم : 16 3 14 15 17 18 17 13 14 16 يكون المدى $R = 18 - 3 = 15$

16 15 14 15 17 18 17 13 14 16 يكون المدى $R = 18 - 15 = 3$

أي أن المدى واحد للمجموعتين في حين يبدو للعين المجردة أن هناك تشتت للبيانات أكبر في المجموعة الأولى عنه في المجموعة الثانية ، مما يعني أن المدى هنا لا يظهر هذا الفارق.

لذا يُعد المدى مقياساً للتشتت لكنه غير جيد في كثير من الأحيان

- لا يمكن تحديده في حالة التوزيعات التكرارية المفتوحة

الفئة	العمر x
الأولى	$x < 6$
الثانية	$6 \leq x < 12$
الثالثة	$12 \leq x < 15$
الرابعة	$15 \leq x < 18$

مفتوح من أسفل

الفئة	العمر x
الأولى	$6 \leq x < 12$
الثانية	$12 \leq x < 15$
الثالثة	$15 \leq x < 18$
الرابعة	$x \geq 18$

مفتوح من أعلى

الفئة	العمر x
الأولى	$x < 6$
الثانية	$6 \leq x < 12$
الثالثة	$12 \leq x < 15$
الرابعة	$x \geq 15$

مفتوح من الطرفين

لا يمكن تحديد مدى البيانات

- لا يدخل في حسابه جميع البيانات

مثال :

البيانات التالية لدرجات ذكاء مجموعتين من الأطفال أوجد المدى وقارن بين المجموعتين:

	القيمة الكبرى للذكاء	القيمة الصغرى للذكاء	متوسط الذكاء
المجموعة A	112	90	105
المجموعة B	140	75	120

A المدى $= 140 - 75 = 65$ درجة

B المدى $= 112 - 90 = 22$ درجة

مثال

البيانات التالية لدرجات الطلاب في مقرر الإحصاء الاجتماعي. احسب المدي لدرجات الطلاب؟

فئات الدرجات	50 —	58 —	66 —	74 —	82 —	90 — 98
عدد الطلاب	3	10	24	40	15	8

المدى $= 98 - 50 = 48$

ثانياً : الانحراف المتوسط [أو متوسط الانحرافات] M.D

يُعرف الانحراف المتوسط (أو متوسط الانحرافات) [وسنرمز له بالرمز $M.D$] على أنه متوسط القيم المطلقة للانحرافات عن قيمة متوسطة للبيانات [عادةً تكون الوسط الحسابي أو الوسيط] .

فإذا اعتبرنا أن القيمة المتوسطة للبيانات هي الوسط الحسابي ، فإن الانحراف المتوسط لمجموعة من البيانات عددها n يُعطى بـ :

$$M.D=\frac{\sum|x-\bar{x}|}{n} \quad \text{أو} \quad M.D=\frac{\sum|d|}{n}$$

$d=x-\bar{x}$ الانحراف عن الوسط

n عدد القيم

ملحوظة هامة : القيمة المطلقة لأي عدد y هي القيمة العددية له دون إشارة ، ونرمز له بنفس الرمز y لكن بين خطين رأسيين | | ، أي نكتب القيمة المطلقة لـ y على الصورة $|y|$. فمثلاً :

$|3| = 3 \quad , \quad |-3| = 3 \quad , \quad |2.5| = 2.5 \quad , \quad |-3.25| = 3.25$

وهكذا .

حيث $d=x-\bar{x}$ هي انحراف القيمة x عن الوسط الحسابي ، $|d|$ هي القيمة المطلقة للانحراف d .

احسب الوسط الحسابي ← ثم احسب الانحرافات عن الوسط ← ثم احسب القيم المطلقة للانحرافات ← ثم احسب متوسط القيم المطلقة

$$\bar{x}=\frac{15+13+3+5+18+12+6+7+3+15}{10}=9.7$$

15	13	3	5	18	12	6	7	3	15
-9.7	-9.7	-9.7	-9.7	-9.7	-9.7	-9.7	-9.7	-9.7	-9.7
5.3	3.3	-6.7	-4.7	8.3	2.3	-3.7	-2.7	-6.7	5.3
5.3	3.3	6.7	4.7	8.3	2.3	3.7	2.7	6.7	5.3

لاحظ أن مجموع الانحرافات = صفر

إذن الانحراف المتوسط هو متوسط القيم المطلقة للانحرافات :

$$M.D=\frac{5.3+3.3+6.7+4.7+8.3+2.3+3.7+2.7+6.7+5.3}{10}=\underline{\underline{4.9}}$$

وسطها الحسابي :

$$\frac{16+14+13+17+18+17+15+14+3+16}{10}=14.3$$

16	14	13	17	18	17	15	14	3	16
-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3	-14.3
1.7	-0.3	-1.3	2.7	3.7	2.7	0.7	-0.3	-11.3	1.7
1.7	0.3	1.3	2.7	3.7	2.7	0.7	0.3	11.3	1.7

لاحظ أن مجموع الانحرافات = صفر

$$M.D=\frac{1.7+0.3+1.3+2.7+3.7+2.7+0.7+0.3+11.3+1.7}{10}=\underline{\underline{2.64}}$$

أي أن الانحراف المتوسط للمجموعة الثانية من القيم أقل من الانحراف المتوسط للمجموعة الأولى من القيم مما يعني أن المجموعة الثانية أقل تشتتاً من المجموعة الأولى وهذا الأمر لا يمكن ملاحظته عند استخدام المدى كمقياس للتشتت

ويمكن أن يتم حل السؤال السابق وذلك بتنظيم خطواتنا من خلال جداول كالتالي:

المجموعة الأولى [$n = 10$]

x	$\bar{x}$	$d=x-\bar{x}$	$\lvert d\rvert$
15	9.7	15 – 9.7 = 5.3	5.3
13	9.7	13 – 9.7 = 3.3	3.3
3	9.7	3 – 9.7 = -6.7	6.7
5	9.7	5 – 9.7 = -4.7	4.7
18	9.7	18 – 9.7 = 8.3	8.3
12	9.7	12 – 9.7 = 2.3	2.3
6	9.7	6 – 9.7 = -3.7	3.7
7	9.7	7 – 9.7 = -2.7	2.7
3	9.7	3 – 9.7 = -6.7	6.7
15	9.7	15 – 9.7 = 5.3	5.3
97	97	0	49
$\sum x$		$\sum d$	$\sum \lvert d\rvert$

$$\bar{x}=\frac{\sum x}{n}=\frac{97}{10}=9.7 \qquad M.D=\frac{\sum|d|}{n}=\frac{49}{10}=4.9$$

المجموعة الثانية [$n = 10$]

x	$\bar{x}$	$d=x-\bar{x}$	$\lvert d\rvert$
16	14.3	16 – 14.3 = 1.7	1.7
14	14.3	14 – 14.3 = -0.3	0.3
13	14.3	13 – 14.3 = -1.3	1.3
17	14.3	17 – 14.3 = 2.7	2.7
18	14.3	18 – 14.3 = 3.7	3.7
17	14.3	17 – 14.3 = 2.7	2.7
15	14.3	15 – 14.3 = 0.7	0.7
14	14.3	14 – 14.3 = -0.3	0.3
3	14.3	3 – 14.3 = -11.3	11.3
16	14.3	16 – 14.3 = 1.7	1.7
143	143	0	26.4
$\sum x$		$\sum d$	$\sum \lvert d\rvert$

$$\bar{x}=\frac{\sum x}{n}=\frac{143}{10}=14.3 \qquad M.D=\frac{\sum|d|}{n}=\frac{26.4}{10}=2.64$$

- ## وفي حالة البيانات الكمية المتقطعة ذات التكرارات:

- يمكن تحديد الانحراف المتوسط $M.D$ من العلاقة :

$$M.D=\frac{\sum f\,|d|}{\sum f}$$

أي نضرب القيمة المطلقة لانحراف كل قيمة [عن الوسط] في تكرارها ، ثم نقسم الناتج على مجموع التكرارات

فمثلاً : إذا كان المطلوب حساب الانحراف المتوسط للبيانات المبينة بالجدول التكراري :

التكرار f	المتغير x
20	4
40	5
30	6
10	7

الجدول التكراري

المتغير x	التكرار f	fx	$d=x-\bar{x}$	$\|d\|$	$f\|d\|$
4	20	80	$4-5.3=-1.3$	1.3	$20\times1.3=26$
5	40	200	$5-5.3=-0.3$	0.3	$40\times0.3=12$
6	30	180	$6-5.3=0.7$	0.7	$30\times0.7=21$
7	10	70	$7-5.3=1.7$	1.7	$10\times1.7=17$
	100	530			76
	$\sum f=100$	$\sum fx=530$			$\sum f\|d\|$

$$\bar{x}=\frac{\sum fx}{\sum f}=\frac{530}{100}=5.3$$

$$M.D=\frac{\sum f\,|d|}{\sum f}=\frac{76}{100}=0.76$$

انتبه :

مجموع الانحرافات هنا [والذي يجب أن يساوي صفراً] هو $\sum fd$ وليس $\sum d$

- **وفي حالة البيانات الكمية المتصلة :**

- نستخدم نفس العلاقة السابقة لتحديد الانحراف المتوسط $M.D$ ، أي يكون

$$M.D=\frac{\sum f\,|d|}{\sum f}$$ حيث $$d=x_0-\bar{x}$$ ، x_0 تمثل مراكز الفئات

الفئة	المتغير x	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	$\|d\|$	$f\,\|d\|$
الأولى	$50 \le x < 60$	6	55	330	$55 - 83.75 = -28.75$	28.75	172.5
الثانية	$60 \le x < 70$	9	65	585	$65 - 83.75 = -18.75$	18.75	168.75
الثالثة	$70 \le x < 80$	15	75	1125	$75 - 83.75 = -8.75$	8.75	131.25
الرابعة	$80 \le x < 90$	12	85	1020	$85 - 83.75 = 1.25$	1.25	15
الخامسة	$90 \le x < 100$	9	95	855	$95 - 83.75 = 11.25$	11.25	101.25
السادسة	$100 \le x < 120$	6	110	660	$110 - 83.75 = 26.25$	26.25	157.5
السابعة	$120 \le x < 180$	3	150	450	$150 - 83.75 = 66.25$	66.25	198.75
		60		5025			945
		$\sum f$		$\sum fx_0$			$\sum f\,\|d\|$

$$\bar{x}=\frac{\sum f\,x_0}{\sum f}=\frac{5025}{60}=83.75$$

$$M.D=\frac{\sum f\,|d|}{\sum f}=\frac{945}{60}=15.75$$

ثالثاً : التباين s^2 والانحراف المعياري s

يُعرف متوسط مربعات الانحرافات عن الوسط الحسابي على أنه تباين مجموعة البيانات [ويُرمز له بالرمز s^2] ، ويُعرف الجذر التربيعي للتباين على أنه الانحراف المعياري للبيانات [ويُرمز له بالرمز s] ، أي أن :

$$\text{التباين} = s^2=\frac{\sum d^2}{n}$$

ومنه يكون

$$\text{الانحراف المعياري} = \sqrt{\frac{\sum d^2}{n}}=\sqrt{s^2}=s$$

المجموعة الأولى [$n = 10$]		
x	$d = x - \bar{x}$	d^2
15	15 – 9.7 = 5.3	28.09
13	13 – 9.7 = 3.3	10.89
3	3 – 9.7 = -6.7	44.89
5	5 – 9.7 = -4.7	22.09
18	18 – 9.7 = 8.3	68.89
12	12 – 9.7 = 2.3	5.29
6	6 – 9.7 = -3.7	13.69
7	7 – 9.7 = -2.7	7.29
3	3 – 9.7 = -6.7	44.89
15	15 – 9.7 = 5.3	28.09

$$\bar{x} = \frac{\sum x}{n} = 9.7$$

$$s^2 = \frac{\sum d^2}{n} = \frac{274.1}{10} = 27.41$$

$$s = \sqrt{s^2} = \sqrt{27.41} \cong \underline{\underline{5.24}}$$

المجموعة الثانية [$n = 10$]		
x	$d = x - \bar{x}$	d^2
16	16 – 14.3 = 1.7	2.89
14	14 – 14.3 = -0.3	0.09
13	13 – 14.3 = -1.3	1.69
17	17 – 14.3 = 2.7	7.29
18	18 – 14.3 = 3.7	13.69
17	17 – 14.3 = 2.7	7.29
15	15 – 14.3 = 0.7	0.49
14	14 – 14.3 = -0.3	0.09
3	3 – 14.3 = -11.3	127.69
16	16 – 14.3 = 1.7	2.89

$$\bar{x} = \frac{\sum x}{n} = 14.3$$

$$s^2 = \frac{\sum d^2}{n} = \frac{164.1}{10} = 16.41$$

$$s = \sqrt{s^2} = \sqrt{16.41} \cong \underline{\underline{4.05}}$$

- **وفي حالة البيانات الكمية المتقطعة ذات التكرارات :**
- يمكن تحديد التباين s^2 والانحراف المعياري s من :

$$\text{التباين} = s^2 = \frac{\sum fd^2}{\sum f}$$

ومنه يكون

$$\text{الانحراف المعياري} = s = \sqrt{s^2} = \sqrt{\frac{\sum fd^2}{\sum f}}$$

فمثلاً : إذا كان المطلوب حساب الانحراف المعياري للبيانات المبينة بالجدول التكراري

الجدول التكراري

التكرار f	المتغير x
20	4
40	5
30	6
10	7

الجدول التكراري					
المتغير x	التكرار f	fx	$d = x - \bar{x}$	d^2	fd^2
4	20	80	4 – 5.3 = -1.3	1.69	20 × 1.69 = 33.8
5	40	200	5 – 5.3 = -0.3	0.09	40 × 0.09 = 3.6
6	30	180	6 – 5.3 = 0.7	0.49	30 × 0.49 = 14.7
7	10	70	7 – 5.3 = 1.7	2.89	10 × 2.89 = 28.9
	100	530			81
	$\sum f = 100$	$\sum fx = 530$			$\sum fd^2$

$$\bar{x} = \frac{\sum fx}{\sum f} = \frac{530}{100} = 5.3$$

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{81}{100} = 0.81$$

$$s = \sqrt{s^2} = \sqrt{0.81} = \underline{\underline{0.9}}$$

- **وفي حالة البيانات الكمية المتصلة :**

- تُستخدم نفس العلاقة السابقة لتحديد الانحراف المعياري s:

$$\boxed{\text{التباين} = s^2 = \frac{\sum fd^2}{\sum f}} \xrightarrow[\text{يكون}]{\text{ومنه}} \boxed{\text{الانحراف المعياري} = s = \sqrt{s^2} = \sqrt{\frac{\sum fd^2}{\sum f}}}$$

حيث $\boxed{d = x_0 - \bar{x}}$

أي أنه عند حساب الانحرافات نعتبر أن مركز أي فئة يمثل جميع القيم الموجودة في تلك الفئة:

الفئة	المتغير x	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	d^2	fd^2
الأولى	$0 \le x < 20$	4	10	40	$10 - 31.7 = -21.7$	470.89	$4 \times 470.89 = 1883.56$
الثانية	$20 \le x < 30$	16	25	400	$25 - 31.7 = -6.7$	44.89	$16 \times 44.89 = 718.24$
الثالثة	$30 \le x < 35$	12	32.5	390	$32.5 - 31.7 = 0.8$	0.64	$12 \times 0.64 = 7.68$
الرابعة	$35 \le x < 40$	10	37.5	375	$37.5 - 31.7 = 5.8$	33.64	$10 \times 33.64 = 336.4$
الخامسة	$40 \le x < 50$	6	45	270	$45 - 31.7 = 13.3$	176.89	$6 \times 176.89 = 1061.34$
السادسة	$50 \le x < 60$	2	55	110	$55 - 31.7 = 23.3$	542.89	$2 \times 542.89 = 1085.78$
		50		1585			5093
		$\sum f$		$\sum fx_0$			$\sum fd^2$

$$\therefore \quad \bar{x} = \frac{\sum f x_0}{\sum f} = \frac{1585}{50} = \underline{\underline{31.7}}$$

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{5093}{50} = 101.86$$

$$s = \sqrt{s^2} = \sqrt{101.86} \cong \underline{\underline{10.09}}$$

<u>وبنفس الأسلوب يمكن التعامل مع المثال السابق لحساب التباين والانحراف المعياري</u>

الفئة	المتغير x	التكرار f	المركز x_0	fx_0	$d = x_0 - \bar{x}$	d^2	$f \times d^2$
الأولى	$50 \le x < 60$	6	55	330	$55 - 83.75 = -28.75$	826.56	4959.38
الثانية	$60 \le x < 70$	9	65	585	$65 - 83.75 = -18.75$	351.56	3164.04
الثالثة	$70 \le x < 80$	15	75	1125	$75 - 83.75 = -8.75$	76.56	1148.4
الرابعة	$80 \le x < 90$	12	85	1020	$85 - 83.75 = 1.25$	1.56	18.72
الخامسة	$90 \le x < 100$	9	95	855	$95 - 83.75 = 11.25$	126.56	1139.04
السادسة	$100 \le x < 120$	6	110	660	$110 - 83.75 = 26.25$	689.06	4134.36
السابعة	$120 \le x < 180$	3	150	450	$150 - 83.75 = 66.25$	4389.06	13167.18
		60		5025			27731.12
		$\sum f$		$\sum fx_0$			$\sum fd^2$

$$\bar{x} = \frac{\sum f x_0}{\sum f} = \frac{5025}{60} = \underline{\underline{83.75}}$$

$$s^2 = \frac{\sum fd^2}{\sum f} = \frac{27731.12}{60} \cong 462.19$$

$$s = \sqrt{s^2} = \sqrt{462.19} \cong \underline{\underline{21.5}}$$

من السابق يتضح أن كلاً من الانحراف المتوسط والانحراف المعياري يعتمدا تماماً في حساباتهما على الوسط الحسابي ، وبالتالي فلهما نفس مزايا وعيوب الوسط الحسابي .
أي :

المزايا :

- من السهل حسابهما
- يأخذ في الاعتبار جميع البيانات
- لا يحتاجا لترتيب معين للبيانات

العيوب :

- يتأثرا بشدة بالقيم المتطرفة
- لا يمكن إيجادهما بالرسم (بيانياً)
- لا يمكن حسابهما للتوزيعات التكرارية المفتوحة

- ويمكن تلخيص كل ما يخص الوسط الحسابي والانحراف المتوسط والانحراف المعياري في الآتي :

- للقيم المفردة :

مربع الانحرافات	القيم المطلقة للانحرافات	الانحرافات عن الوسط	قيم عددها n
d^2	$\|d\|$	$d = x - \bar{x}$	x
...	...	...	...
...	...	...	...
$\sum d^2$	$\sum \|d\|$		$\sum x$

الوسط الحسابي = $\bar{x} = \dfrac{\sum x}{n}$

الانحراف المتوسط = $M.D = \dfrac{\sum |d|}{n}$

$s^2 = \dfrac{\sum d^2}{n} \longrightarrow s = \sqrt{s^2}$

- **ولتوزيع تكراري :**

القيم	التكرار		الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات		
x	f	fx	$d = x - \bar{x}$	$\lvert d\rvert$	d^2	$f\lvert d\rvert$	fd^2
...	...	...	...	...	...	...	...
...	...	...	...	...	...	...	...
	$\sum f$	$\sum fx$				$\sum f\lvert d\rvert$	$\sum fd^2$

$$\bar{x} = \frac{\sum fx}{\sum f}$$

$$M.D = \frac{\sum f\lvert d\rvert}{\sum f}$$

$$\longrightarrow s = \sqrt{s^2}$$

- **وللبيانات المتصلة :**

الفئات	التكرار			الانحرافات عن الوسط	القيم المطلقة للانحرافات	مربع الانحرافات		
x	f	مراكز الفئات	$f x_0$	$d = x_0 - \bar{x}$	$\lvert d\rvert$	d^2	$f\lvert d\rvert$	fd^2
...	...	x_0	...	...	...	...	...	...
...	...	...	...	...	...	...	...	...
	$\sum f$	...	$\sum fx$				$\sum f\lvert d\rvert$	$\sum fd^2$

خاصيتان هامتان للانحراف المتوسط والانحراف المعياري :

الخاصية الأولى : إضافة عدد ثابت c لكل قيمة من قيم البيانات لا يؤثر على قيمة الانحرافين المتوسط والمعياري .

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم

الخاصية الثانية : ضرب كل قيمة من قيم البيانات في عدد ثابت c يجعل :

الانحراف المتوسط (أو المعياري) الجديد = الانحراف المتوسط (أو المعياري) القديم × القيمة المطلقة للثابت c

فمثلاً، لو كانت لدينا البيانات التالية والتي توضح درجات مجموعة من الطلاب كالتالي :

الدرجات الأصلية			
x	d	$\|d\|$	d^2
9	1	1	1
2	-6	6	36
7	-1	1	1
12	4	4	16
10	2	2	4
40		14	58

$$\bar{x}=\frac{\sum x}{n}=\frac{40}{5}=\underline{\underline{8}}$$

$$M.D=\frac{\sum|d|}{n}=\frac{14}{5}=\underline{\underline{2.8}}$$

$$s^2=\frac{\sum d^2}{n}=\frac{58}{5}=11.6$$

$$s=\sqrt{11.6}\cong\underline{\underline{3.4}}$$

بعد إضافة 5 لكل درجة			
x	d	$\|d\|$	d^2
14	1	1	1
7	-6	6	36
12	-1	1	1
17	4	4	16
15	2	2	4
65		14	58

$$\bar{x}=\frac{\sum x}{n}=\frac{65}{5}=\underline{\underline{13}}$$

$$M.D=\frac{\sum|d|}{n}=\frac{14}{5}=\underline{\underline{2.8}}$$

$$s^2=\frac{\sum d^2}{n}=\frac{58}{5}=11.6$$

$$s=\sqrt{11.6}\cong\underline{\underline{3.4}}$$

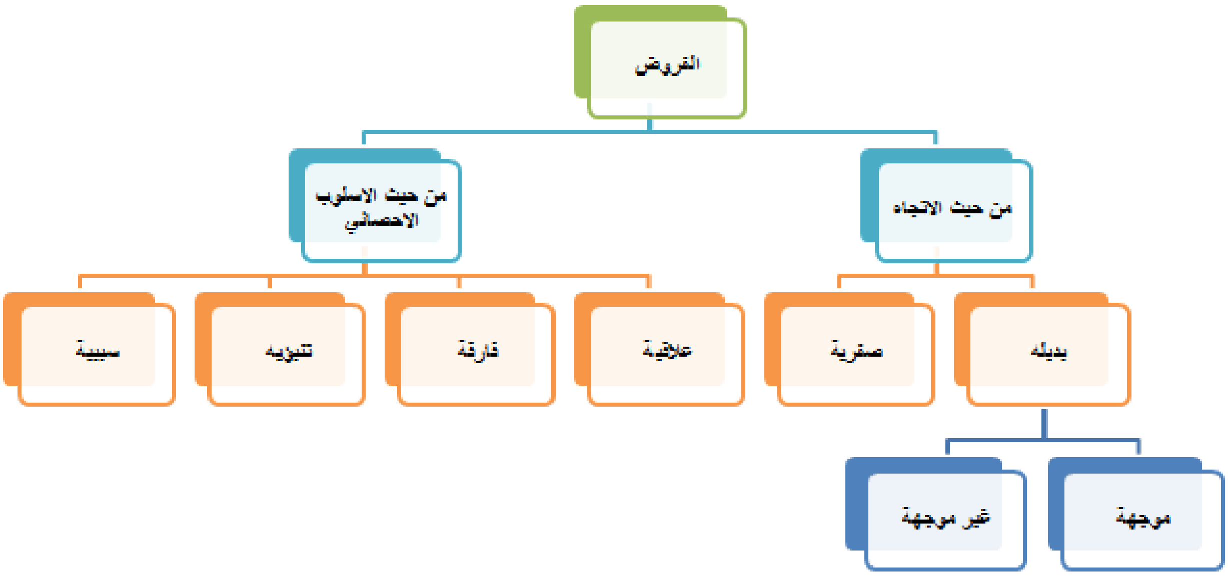
الفروض الإحصائية

يعرف الفرض بأنه إجابة متوقعة لسؤال من الأسئلة التي تراود ذهن الباحث أو المهتم، وهذه الإجابة لا تكون نهائية وإنما خاضعة للدراسة والتحقق من مدى صحتها فإما أن تكون الإجابة صحيحة وإما أن تكون الإجابة خاطئة.

وتوقع الإجابة من جانب الباحث لا يتم من فراغ وإنما بناءً على خلفية نظرية متعلقة بهذا السؤال ونتائج دراسات سابقة حوله.

فمثلاً: يراود ذهن الباحث سؤال مضمونه: ما طبيعة العلاقة بين حب الاستطلاع والقدرة الابتكارية لدى طلاب قسم علم الاجتماع؟ وبناء على الخلفية النظرية ونتائج الدراسات السابقة المرتبطة بطبيعة العلاقة بين المتغيرين يصيغ الباحث إجابة متوقعة لهذا السؤال وهي تمثل إحدى فروض بحثه وتكون صياغة الفرض كالتالي:

- توجد علاقة بين حب الاستطلاع والابتكارية
- لا توجد علاقة بين حب الاستطلاع والابتكارية
- الفرض هو اقتراح لقضية معينة وبالتالي فإن قرار قبولنا هذا الاقتراح كاقتراح صحيح أو رفضنا إياه كاقتراح خاطئ لا بد أن يؤجل حتى نجمع دليل يؤكد قبوله أو رفضه.



- **الفرضية الصفرية (فرضية العدم)** H_0 (Null Hypothesis) :

هي الفرضية حول معلمة المجتمع التي نجري اختبار عليها باستخدام بيانات من عينة والتي تشير أن الفرق بين معلمة المجتمع والإحصائي من العينة ناتج عن الصدفة ولا فرق حقيقي بينهما. وهي الفرضية التي ننطلق منها ونرفضها عندما تتوفر دلائل على عدم صحتها، وخلاف ذلك نقبلها وتعني كلمة Nul انه لا يوجد فرق بين معلمة المجتمع والقيمة المدعاة (إحصائية العينة).

- **الفرضية البديلة** Alternative Hypothesis (H_a) :

- هي الفرضية التي يضعها الباحث كبديل عن فرضية العدم و نقبلها عندما نرفض فرضية العدم باعتبارها ليست صحيحة بناء على المعلومات المستقاة من العينة.

- وفي اختبار الفروض يمكن أن نرتكب نوعين من الخطأ:

- <u>الخطأ من النوع الأول Type I error:</u> الخطأ من النوع الأول هو "رفض الفرض العدمي بينما هو صحيح ". أي أنه على الرغم من أن الفرض العدمي في الواقع صحيح وكان من الواجب قبوله فقد تم أخذ قرار خاطئ برفضه. وباختصار شديد فإن الخطأ من النوع الأول هو : " رفض فرض صحيح". ويرمز له بالرمز α .

- <u>الخطأ من النوع الثاني Type II error:</u> وفي المقابل فإن الخطأ من النوع الثاني يعني " قبول الفرض العدمي بينما هو خاطئ " أي أنه على الرغم من أن الفرض العدمي خاطئ وكان من الواجب رفضه فقد تم أخذ قرار خاطئ بقبوله وباختصار شديد فإن الخطأ من النوع الثاني هو " قبول فرض خاطئ ". ويرمز له بالرمز β .

- <u>ويمكن أن نمثل ذلك في الجدول التالي:</u>

القرار \ الفرضية	صحيحة (H_0)	خاطئة (H_a)
قبول (H_0)	صواب	خطأ ٢ بيتا (B)
رفض (H_0)	خطأ ١ ألفا (a)	صواب

١. فرضية صحيحة نتائج العينة تؤيد صحتها. (قبول صواب)

١. فرضية صحيحة نتائج العينة غير مؤيدة لصحتها. (رفض صواب)
وهذا يعطينا خطأ من النوع الأول ألفا (a)

١. فرضية خاطئة نتائج تؤيد صحتها (قبول خطأ) وهذا يعطينا خطأ من النوع الثاني بيتا (B) ويمكن أن يقلل بزيادة حجم العينة

١. فرضية خاطئة نتائج غيرمؤيدة صحتها (رفض خطأ)

الفروض البحثية:

هي الفروض التي يصيغها الباحث في بحثه بناءً على خلفيته النظرية ونتائج الدراسات السابقة.

١. الفروض العلاقية:

أ. الفرض البديل العلاقي غير الموجه:

توجد علاقة دالة إحصائياً بين الاتجاه نحو الدراسة والبيئة الدراسية

ب. الفرض البديل العلاقي الموجه:

توجد علاقة ايجابية دالة إحصائياً بين الاتجاه نحو الدراسة والبيئة الدراسية

ج. الفرض الصفري العلاقي:

لا توجد علاقة دالة إحصائياً بين نحو الدراسة والبيئة الدراسية

٢. الفروض الفارقة:

أ. الفرض البديل الفارق غير الموجه:

توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء الوجداني

ب. الفرض البديل الفارق الموجه:

توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء الوجداني لصالح الذكور

ج. الفرض الصفري الفارق:

لا توجد فروق ذات دلالة إحصائية بين متوسطي درجات الذكور والإناث في الذكاء الوجداني

٣. الفروض التنبؤية:

أ. الفرض البديل التنبؤي غير الموجه:

يمكن التنبؤ من المتغيرات المستقلة (الدافعية، وحب الاستطلاع، والقلق) بالمتغير التابع (التحصيل الدراسي) لدى طلاب جامعة الملك فيصل

ب. الفرض البديل التنبؤي الموجه:

يمكن التنبؤ من المتغيرات المستقلة (الدافعية كمنبئ موجب، وحب الاستطلاع كمنبئ موجب، والقلق كمنبئ سالب) بالمتغير التابع (التحصيل الدراسي) لدى طلاب جامعة الملك فيصل

ج. الفرض الصفري التنبؤي :

لا يمكن التنبؤ من المتغيرات المستقلة (الدافعية، وحب الاستطلاع، والقلق) بالمتغير التابع (التحصيل الدراسي) لدى طلاب جامعة الملك فيصل

٤. الفروض السببية:

أ. الفرض البديل السببي غير الموجه:

يمكن التوصل إلى نموذج سببي يفسر العلاقة بين المتغيرات المستقلة (المعاملة الوالدية، والذكاء، والضغوط النفسية، والاتجاه نحو الدراسة) والمتغير التابع (مستوى الطموح) لدى طلاب جامعة الملك فيصل

ب. الفرض البديل السببي الموجه:

يمكن التوصل إلى نموذج سببي يفسر العلاقة بين المتغيرات المستقلة (المعاملة الوالدية «تأثير موجب»، والذكاء «تأثير موجب»، والضغوط النفسية تأثير سالب»، والاتجاه نحو الدراسة «تأثير موجب») والمتغير التابع (مستوى الطموح) لدى طلاب جامعة الملك فيصل

ج. الفرض الصفري السببي:

لا يمكن التوصل إلى نموذج سببي يفسر العلاقة بين المتغيرات المستقلة (المعاملة الوالدية، والذكاء، والضغوط النفسية، والاتجاه نحو الدراسة) والمتغير التابع (مستوى الطموح) لدى طلاب جامعة الملك فيصل

الفروض الإحصائية:

ما الفرق بين الفروض البحثية والفروض الاحصائية؟

الفروض البحثية

هي الفروض التي يصيغها الباحث بنفسه في ضوء اطلاعه على الخلفية النظرية ونتائج الدراسات السابقة، وبناء على اطلاعه يحدد اتجاه الفرض هل هو فرض بديل موجه أم فرض بديل غير موجه أم فرض صفري.

أما الفروض الإحصائية

فتهدف إلى تفسير نتيجة معالجة الأسلوب الإحصائي للفرض البحثي، والذي بناء عليه نتقبل الفرض البحثي أو نرفضه، وبالتالي فالذي يجعلنا نقبل الفرض البحثي ليس الأسلوب الإحصائي فقط ولكن الفرض الإحصائي المرتبط به.