كلية العلوم

القسم : **علم الحياة**

السنة : **الاولى**

المادة : احصاء حيوي

المحاضرة : ٧+٨ /نظري/د.نبيل

5-5 العلاقة بين مقاييس النزعة المركزية ومقاييس التشتت

1- معامل الالتواء

2- معامل الاختلاف

3- الدرجة المعيارية

4- التفرطح.

5-5-1 معامل الالتواء $\propto$ Skewness

هو معامل الالتواء "لبيرسون" ويعطى بالعلاقة: $\propto = X\,\frac{3(\bar{X}-Med)}{S}$

ويمكن من خلال الإشارة التي يأخذها هذا المعامل الحكم على شكل الالتواء كما يلي:

- إذا كان (الوسط الحسابي = الوسيط) كان قيمة المعامل ($\propto = 0$) ويدل على أن المنحني التوزيع التكراري متماثل.
- إذا كان (الوسط الحسابي > الوسيط) كان قيمة المعامل ($\propto > 0$) ويدل على أن المنحني التوزيع التكراري ملتوي جهة اليمين.
- إذا كان (الوسط الحسابي < الوسيط) كان قيمة المعامل ($\propto < 0$) ويدل على أن المنحني التوزيع التكراري ملتوي جهة اليسار.

مثال 5-5-1` : البيانات التالية تمثل درجات 8 طلاب في مقرر الإحصاء :

66 85 52 78 80 91 74 58

المطلوب :حساب معامل الالتواء.

الحل: بالقيام بالحسابات اللازمة نجد أن:

الوسط الحسابي $\bar{X} = 73$ ، الوسيط : Med=76، الانحراف المعياري : S=13.406

فيكون معامل الالتواء $\propto = \frac{3(\bar{X}-Med)}{S} = \frac{3(73-76)}{13.406} = -0.67$

إذاً منحني توزيع درجات الطلاب ملتوي جهة اليسار.

2-5-5 معامل الاختلاف Variation Coefficient

معامل الاختلاف عبارة عن طريقة تستخدم للمقارنة بين تشتت مجموعات مختلفة ، فهو أحد المقاييس المستخدمة لقياس درجة التشتت، وفيه يحسب قيمة التشتت كنسبة مئوية من قيمة مقياس النزعة المركزية، ومن ثم يفضل استخدام معامل الاختلاف عند مقارنة درجة تشتت بيانات مجموعتين أو أكثر مختلفة لها وحدات قياس مختلفة بدلا من الانحراف المعياري لأن معامل الاختلاف يعتمد على التغيرات النسبية في القيم عند قياس النزعة المركزية، بينما يعتمد الانحراف المعياري على التغيرات المطلقة للقيم ، فعند مقارنة درجة التشتت بيانات الأطوال بالسنتيمتر وبيانات الأوزان بالكيلوغرام لا يمكن الاعتماد على الانحراف المعياري في هذه المقارنة وإنما يستخدم معامل الاختلاف ومن ثم يطلق عليه الاختلاف النسبي ويحسب معامل الاختلاف النسبي بتطبيق العلاقة التالية:

$$C.V = \frac{S}{\overline{X}} \times 100$$

كلما كان معامل الاختلاف لمجموعة من البيانات x أقل من الاختلاف لمجموعة أخرى y فهذا يعني أن بيانات المجموعة x أكثر تجانساً من المجموعة y ويكون درجة تشتت بيانات x أقل من درجة تشتت بيانات المجموعة y

مثال 2-5-5: لنفرض أننا نريد المقارنة بين تحصيل طالبين كل منهما في شعبة دراسية في الجدول التالي:

أي من الطالبين أفضل في تحصيله بالنسبة لمستوى شعبته

المقاييس	طالب الشعبة (1)	طالب الشعبة (2)
درجة الطالب	82	60
الوسط الحسابي	70	54
الانحراف المعياري	10	4

الحل: معامل الاختلاف لطالب نسبي الشعبة (1)هو: V.C=10/70=14.28%

معامل الاختلاف النسبي لطالب الشعبة (2)هو : V.C =4/54=7.41%

هذا يعني أن تحصيل طالب الشعبة (2)أفضل بالرغم من أن درجته الخام أقل من درجة طالب الشعبة (1) .

مثال 5-5-3 :تم اختيار مجموعتين من الأغنام النامية في أحد المزارع وتم استخدام عليقة معينة لتسمين المجموعة الأولى بينما تم استخدام عليقة أخرى لتسمين المجموعة الثانية وبعد فترة تم جمع بيانات عن الأوزان المجموعتين بالكيلوجرام وتم الحصول على المقاييس التالية: والمطلوب مقارنة تشتت المجوعتين .

المقاييس	**المجموعة الأولى**	**المجموعة الثانية**
الوسط الحسابي	173	198
الانحراف المعياري	23	25

الحل:

معامل الاختلاف النسبي للمجموعة الأولى: V.C =(23 . 173)100=13.3%

معامل الاختلاف النسبي للمجموعة الثانية: V.C =(25 . 198)100=12.37%

يلاحظ أن درجة تشتت أوزان المجموعة الثانية أقل من درجة تشتت أوزان المجموعة الأولى

5-5-3:الدرجة المعيارية standardized degree

إن مقاييس التشتت تمكننا من المقارنة بين تشتت مجموعتين مختلفتين لكن إ أردنا أن نقارن بين قيمتين في المجموعة فإذا فرضنا مثلا أن طالبا حصل على 90 درجة في مادة علم الاجتماع وعلى 80 درجة في مادة الإحصاء فليس من المنطقي أن نقارن بين مستوى الطالب في كل المادتين على اساس الدرجة التي حصل عليها بل لا بد مقارنة الك بالنسبة لتوزيع درجات الطلاب في كل مادة ولإجراء هذه المقارنة يجب أن نقارن بين موضع كل من هاتين القيمتين على التوزيع الخاص بهما أي معرفة مدى بعد الدرجة عن الوسط الحسابي بدلالة وحدات الانحراف المعياري ويتم ذلك بطرح الوسط الحسابي لكل توزيع من القيمة الخاصة به وقسمة الناتج على الانحراف المعياري فنحصل على الدرجة المعيارية (ونرمز لها بالرمزz) أي نحسب الدرجة المعيارية من العلاقة التالية: $z=\frac{X-\bar{X}}{s}$

وتستخدم الدرجات المعيارية في مقارنة مستوى أداء فرد معين بمستوى أداء المجموعة التي ينتمي إليها، مقارنة قيمتين أو أكثر مختلفة من حيث وحدات القياس. مقارنة مدى بعد قيمة معينة عن الوسط الحسابي بدلالة وحدات الانحراف المعياري أي توضيح مركز قيمة معينة بالنسبة للمجموعة التي تقع فيها هذه القيمة.

مثال 5-5-5 لنعود إلى مثال السابق:

بفرض طالب حصل على 90 في مادة علم الاجتماع وعلى 80 درجة من مادة الإحصاء وبفرض أن متوسط درجات مادة علم الاجتماع هو 80 درجة والانحراف المعياري هو 5 بينما كان متوسط درجات مادة الاحصاء هو 75 درجة والانحراف المعياري هو 3 والمطلوب: تحديد مستوى الطالب وذلك بمعرفة مدى بعد الدرجة عن الوسط الحسابي بدلالة وحدات الانحراف المعياري.

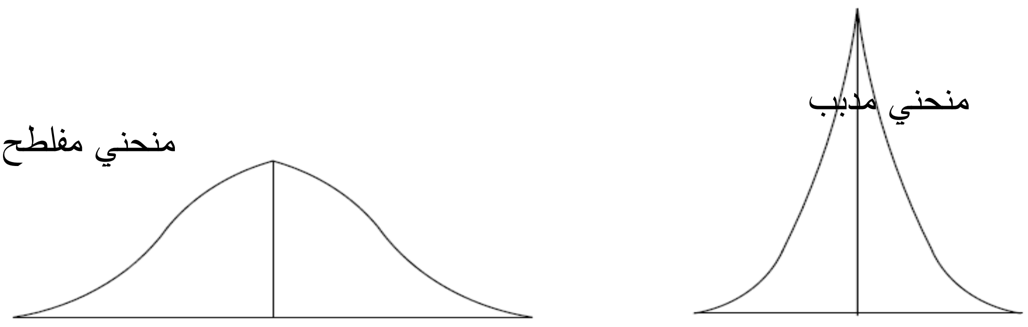
الحل:

الدرجة المعيارية لدرجة الطالب في علم الاجتماع هي: $Z=(90-80)/5=2$ أي أن درجة الطالب في علم الاجتماع تبعد عن المتوسط بقدار 2 درجة معيارية

الدرجة المعيارية لدرجة الطالب في مادة الاحصاء هي: $Z=(80-75)/3=1.67$ أي أن درجة الطالب في مادة الاحصاء تبعد عن المتوسط بقدار 1067 درجة معيارية. من ذلك نجد أن طالب حصل على درجة في علم الاجتماع أبعد عن المتوسط من درجته في الاحصاء فهذا يعني أن مستواه في مادة الاجتماع أعلى من مستواه في الاحصاء على الرغم من ارتفاع درجته فيها.

5-5-4: التفلطح Kurtosis:

عند التمثيل التوزيع التكراري في شكل منحنى تكراري قد يكون هذا منحنى منبسط أو مدبب فعندما يتركز عدد أكبر من القيم بالقرب من منتصف المنحنى ويقل بالقرب من المنتصف يكون المنحنى مفرطحاً او منبسطاً ويظهر في الشكل التالي:



ويتم حساب معامل التفلطح (k) بتطبيق المعادلة التالية : $K=\frac{\frac{1}{n}\sum(X-\overline{X})^4}{S^4}$ معامل التفلطح في التوزيع الطبيعي يساوي 3 ومن ثم يمكن وصف منحني التوزيع من حيث التفلطح و التدبب كما يلي: إذا كان K=3 كان منحني التوزيع معتدلاً

إذا كان k<3 كان منحني التوزيع مدبباً

إذا كان K>3 كان منحني التوزيع منبسطاً

مثال 5-5-7: أوجد شكل منحني التوزيع لدرجات الطلاب المعطاة في المثال 5-5-1:

الحل:

X	$(X-\overline{X})$	$(X-\overline{X})^2$	$(X-\overline{X})^4$
66	−7	49	2401
85	12	144	20736
52	−21	441	194481
78	5	25	625
80	7	49	2401
91	18	324	104976
74	1	1	1
58	−15	225	50625
584	0	1258	376246

من هذه البيانات نجد :

$$S=\sqrt{\frac{\sum(X-\overline{X})^2}{n-1}}=\sqrt{\frac{1258}{7}}=13.406$$

$$\frac{1}{n}\sum(X-\overline{X})^4=\frac{376246}{8}=47030.75$$

بالتعويض في العلاقة 5-5-4 نجد أن معامل التفرطح هو : $K=\frac{47030.75}{32299.58}$ بما أن $k<3$ فإن شكل توزيع بيانات الدرجات مفرطح .

5-6 مقاييس الارتباط والانحدار الخطي :

1- الارتباط الخطي البسيط Simple correlation

2- معامل ارتباط الرتب لبيرسون Ponears

3- معامل ارتباط الرتب سبيرمان Spearman

4- الانحدار الخطي البسيط Simple Regession

5- نموذج الانحدار الخطي

تستخدم مقاييس الارتباط والانحدار الخطي في تحديد درجة العلاقة بين متغيرات المختلفة

تهتم هذه المقاييس بدراسة وتحليل العلاقة بين متغيرين وذلك باستخدام بعض طرق التحليل الإحصائي مثل تحليل الارتباط والانحدار الخطي البسيط فإذا طان اهتمام الباحث وهدفه هو دراسة وتحديد نوع قوة العلاقة بين متغيرين استخدام لذلك استخدام لذلك أسلوب تحليل الارتباط وإذا كان اهتمامه بدراسة أثر أحد المتغيرين على الاخر استخدم لذلك أسلوب تحليل الانحدار ومن الأمثلة على ذلك:

1- الإنفاق والدخل العائلي

2- سعر السلعة والكمية المطلوبة منها

3- تقديرات الطلاب في مقرر الإحصاء وتقديراتهم في مقرر الرياضيات

4- كمية السماد المستخدمة وكمية الإنتاج من حصول معين تم تسميده بهذا النوع من السماد

5 – عدد مرات ممارسة نوع معين من الرياضة البدنية ومستوى الكولسترول في الدم وزن الجسم وضغط الدم

5-6-1 الارتباط الخطي Simple correlation:

هدف تحليل الارتباط الخطي البسيط هو تحديد نوع وقوة العلاقة بين متغيرين ويرمز له غي حالة المجتمع بالرمز وفي حالة العينة بالرمز r أننا في كثير من النواحي التطبيقية نتعامل مع بيانات عينة مسحوبة من المجتمع سوف نهتم بحساب معامل الارتباط في العينة r كتقدير لمعامل ارتباط في المجتمع ومن التحديد السابق للغرض من معامل الارتباط نجد أنه يركز على نقطتين هما:

- نوع العلاقة: وتأخذ ثلاث أنواع حسب إشارة معامل الارتباط كما يلي:

1 – إذا كانت إشارة معامل الارتباط سالبة $(r<0)$ توجد علاقة عكسية بين المتغيرين بمعنى أن زيادة أحد المتغيرين يصاحبه انخفاض في المتغير الثاني والعكس.

2 – إذا كانت إشارة معامل الارتباط موجبة$(r>0)$ توجد علاقة طردية المتغيرين بمعنى أن زيادة أحد المتغيرين يصاحبه زيادة في المتغير الثاني والعكس.

3 – إذا كان معامل ارتباط قيمته صفراً (0) فذلك يدل على انعدام العلاقة بين المتغيرين.

- **قوة العلاقة:** ويمكن الحكم على قوة العلاقة من حيث درجة قربها أو بعدها عن $(1,-1)$ حيث أن قيمة معامل الارتباط تقع في المدى $(-1<r<1)$.

5-6-2: معامل ارتباط الرتب لبيرسون Person:

في حالة جمع البيانات عن متغيرين كميين (x, y) يمكن قياس الارتباط بينهما باستخدام طريقة بيرسون. ومن الأمثلة على ذلك: قياس العلاقة بين الوزن والطول والعلاقة بين الإنتاج والتكلفة والعلاقة بين الإنفاق الاستهلاكي والدخل والعلاقة بين الدرجة التي حصل عليها الطالب وعدد ساعات الدراسة وهكذا الأمثلة على ذلك كثيرة

ولحساب معامل الارتباط في العينة نستخدم صيغة "بيرسون" التالية:

$$r = \frac{S_{xy}}{S_x\, S_y} = \frac{\frac{\sum(X-\bar{X})(Y-\bar{Y})}{n-1}}{\sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}\sqrt{\frac{\sum(Y-\bar{Y})^2}{n-1}}}$$

حيث:

$S_{xy} = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{n-1}$ هو التباين بين (x ,y)

$S_x = \sqrt{\frac{\sum(X-\bar{X})^2}{n-1}}$ هو الانحراف المعياري لقيم X

$S_y = \sqrt{\frac{\sum(Y-\bar{Y})^2}{n-1}}$ هو الانحراف المعياري لقيم Y.

ويمكن اختصار الصيغة السابقة لتصبح على النحو التالي:

$$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2}\sqrt{\sum(Y-\bar{Y})^2}}$$

مثال5-6-1: فيما يلي المساحة المزروعة بالأعلاف الخضراء بالألف هكتار وإجمالي إنتاج اللحوم بالألف طن خلال فترة من 1995 حتى 2002

والمطلوب: حساب معامل الارتباط بين المساحة والكمية وما هو مدلوله

السنة	1995	1996	1997	1998	1999	2000	2001	2002

المساحة	305	313	297	289	233	214	240	217
الكمية	592	603	662	607	635	699	719	747

الحل: بفرض أن x هي المساحة المزروعة و y هي الكمية ولحساب الارتباط بين (x، y) نطبق العلاقة وذلك كما يلي:

- نحسب الوسط الحسابي لكل من المساحة والكمية:

$$\bar{X} = \frac{\sum X}{n} = \frac{2108}{8} = 263.5$$

$$\bar{Y} = \frac{\sum y}{n} = \frac{5264}{8} = 658$$

نحسب المجاميع:

X	Y	$X-\bar{X}$	$(X-\bar{X})^2$	$Y-\bar{Y}$	$(Y-\bar{Y})^2$	$(X-\bar{X})(Y-\bar{Y})$
305	592	41.5	1722.25	−66	4356	−2739
313	603	49.5	2450.25	−55	3025	−2722.5
297	662	33.5	1122.25	4	16	134
289	607	25.5	650.25	−51	2601	−1300.5
233	635	−30.5	930.25	−23	529	701.5
214	699	−49.5	2450.25	41	1681	−2029.5
240	719	−23.5	552.25	61	3721	−1433.5
217	747	−46.5	2162.25	89	7921	−4138.5
2108	5264	0	12040	0	23850	−13528

إذاً يوجد ارتباط عكسي قوي بين المساحة المزروعة وكمية إنتاج اللحوم

$$r = \frac{\sum(X-\bar{X})(Y-\bar{Y})}{\sqrt{\sum(X-\bar{X})^2}\sqrt{\sum(Y-\bar{Y})^2}} = \frac{-13528}{\sqrt{12040}\,\sqrt{23850}} = -\,0.798$$

$$r = \frac{\sum XY - \frac{1}{n}\left(\sum X \sum Y\right)}{\sqrt{\left(\sum X^2 - \frac{\left(\sum X\right)^2}{n}\left(\sum Y^2 - \frac{\left(\sum Y\right)^2}{n}\right)\right)}}$$

يمكن تبسيط العمليات الحسابية للصيغة السابقة لتصبح كما يلي:

3-6-5 معامل ارتباط الرتب سبيرمان Spearman:

إذا كانت الظاهرة محل الدراسة تحتوي على متغيرين وصفيين ترتبين ومثال على ذلك قياس العلاقة بين تقديرات الطلبة في مادتين أو العلاقة بين درجة التفصيل المستهلك لسلعة معينة ومستوى الدخل فإنه يمكن استخدام طريقة "بيرسون" السابقة في حساب معامل ارتباط يعتمد على رتب مستويات المتغيرين كبديل للقيم الأصلية ويطلق على هذا معامل "معامل ارتباط سبيرمان " ويعبر عنه بالمعادلة التالية:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)}$$

حيث إن dهي الفرق بين رتب مستويات المتغير الأول xورتب مستويات الثاني yأي أن d=Rx−Ry

مثال2-5-6: فيما يلي 10 طلاب في مادتي الإحصاء والاقتصاد:

تقديرات إحصاء	أ	ج+	د	د+	ب+	ج+	أ+	ب	ب+	ب+
تقديرات اقتصاد	أ+	د	ج	ج	أ	ب	ب+	ب	ج	ب

والمطلوب:1 − احسب معامل الارتباط بين تقديرات الطلبة في المقررين

2 − وما هو مدلوله

الحل: بإجراء الحسابات اللازمة نحصل على:

$$r_s = 1 - \frac{6\sum d^2}{n(n^2-1)} = 1 - \frac{6(44.5)}{10(99)} = 1 - \frac{267}{990} = 0.7303$$

مدلول معامل ارتباط: بما r=0.7303 فهذا يدل على ارتباط طردي قوي بين تقديرات الطالب في مادة الإحصاء ومادة الاقتصاد.