وزارة التعليم العالي و البحث العلمي
جامعة باجي مختار عنابة
كلية الآداب و العلوم الإنسانية و الاجتماعية
قسم التربية البدنية و الرياضية

مقياس الإحصاء سنة ثانية
الدرس الثاني

إعداد الدكتور: بن شيخ يوسف

قائمة المحتويات

4 – الارتباط و الانحدار
- أنواع الارتباط
- حسب عدد المتغيرات
- حسب نوع العلاقة
- حسب قوة العلاقة

4 . 1 العلاقة بين متغيرين كميين
- لوحة الانتشار
- أنواع لوحة الانتشار
- العلاقات الخطية
- معامل الارتباط الخطي البسيط
- شرح معامل الارتباط الخطي
- معامل (التحديد)
- الدلالة الإحصائية لمعامل التحديد

4 . 2 الارتباط بين متغيرين نوعيين (الرتبي)

4 . 3 الانحدار الخطي

الخطأ المعياري لمعادلة الانحدار أو التنبؤ:

معامل الارتباط الخطي البسيط

بحساب معامل الارتباط الخطي البسيط يمكننا معرفة قوة و نوع العلاقة بين المتغيرين x وy و نرمز له ب r.

و لحساب معامل الارتباط نتبع الخطوات التالية.

- نحسب المقاييس الإحصائية لكل من المتغير x و المتغير y (المتوسط الحسابي و الانحراف المعياري)
- تحويل قيم كل من المتغير x و المتغير y إلى قيم معيارية z
- نحسب حاصل ضرب كل قيمة معيارية للمتغير x و المتغير y ($z_{(y)}.z_{(x)}$)
- نحسب المتوسط الحسابي لحاصل ضرب القيم المعيارية

مثال:

سنقوم بحساب معامل الارتباط للمثال الخاص بالسن و نبض القلب الأقصى

المتغير / الفرد	السن x	نبض القلب الأقصىy	$z_{(x)}$	$z_{(y)}$	$z_{(y)}.z_{(x)}$
1	40	187	0.42 -	0.36 -	0.15
2	36	195	0.82 -	0.99	0.81 -
3	51	180	0.67	1.54 -	1.04 -
4	49	190	0.48	0.14	0.07
5	47	185	0.28	0.70 -	0.19 -
6	51	183	0.67	1.04 -	0.70 -
7	32	195	1.22 -	0.99	1.20 -
8	55	185	1.07	0.70 -	0.75 -
9	55	189	1.07	0.03 -	0.03 -
10	23	201	2.12 -	2.00	4.23 -
11	49	189	0.48	0.03 -	0.01 -
12	52	185	0.77	0.70 -	0.54 -
13	35	195	0.92 -	0.99	0.91 -
المتوسط الحسابي	44.23	189.15	0.00	0.00	**0.78 -**
الانحراف المعياري δ	10.03	5.93	-	-	-

إذن من خلال الجدول يتضح لنا أن قيمة معامل الارتباط هي **r = - 0.78**

و هذا ما يدل على وجود علاقة طردية قوية بين كل من السن و النبض الأقصى للقلب.

هذا المعامل يسمى بمعامل بيرسون للارتباط الخطي البسيط و يمكن حسابه من خلال العلاقة التالية

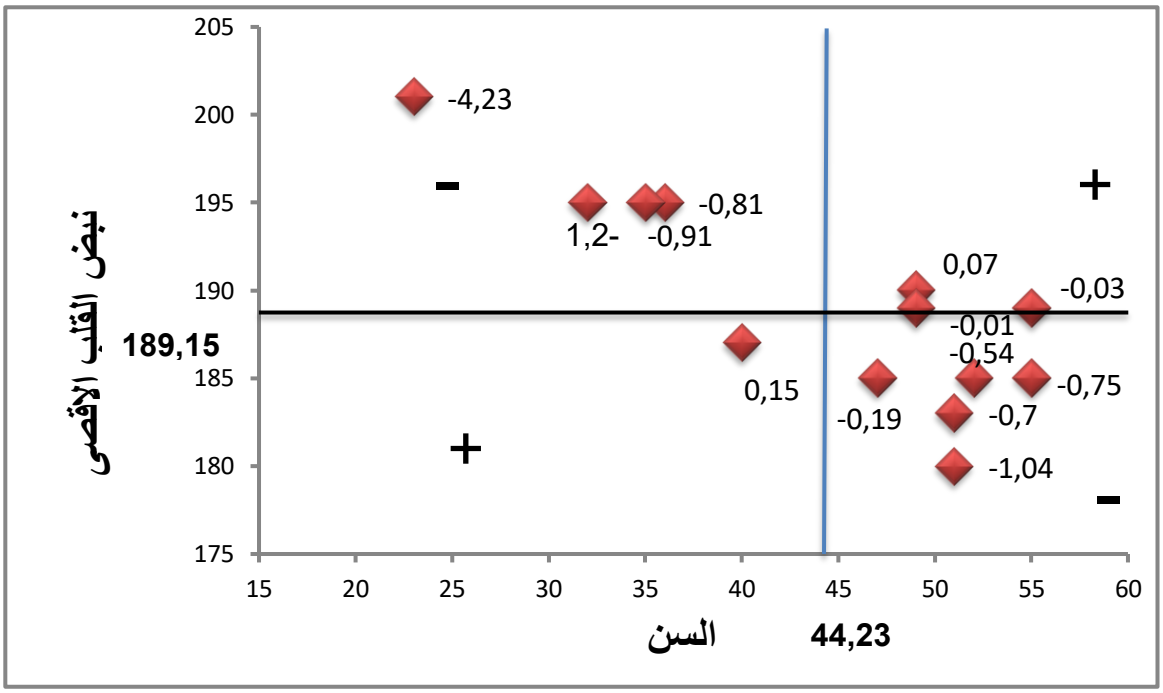
$$r = \frac{\sum(x.y)/n - \bar{x}.\bar{y}}{\delta x \,.\, \delta y}$$

د. يوسف بن شيخ

أستاذ محاضر –أ– بقسم التربية البدنية و الرياضية جامعة باجي مختار عنابة

شرح معامل الارتباط الخطي

الشكل التالي سيوضح لنا مفهوم العلاقة الموجودة بين المتغيرين



المتوسط الحسابي للسن هو 44.23 سنة و المتوسط الحسابي للنبض الأقصى للقلب هو 189.15 سنة هذين المتوسطين يسمحان بتقسيم السحابة النقطية إلى أربعة أقسام.

نلاحظ من خلال الشكل أن معظم النقط تجمعت في الإطارين الأيسر العلوي و الأيمن السفلي حيث نلاحظ تجمع القيم الأكبر من المتوسط الحسابي بالنسبة لنبض القلب الأقصى في الإطار العلوي الأيسر و تجمع القيم الأكبر من المتوسط الحسابي بالنسبة للسن.

إذن نقول أن العلاقة عكسية إذا تجمع عدد كبير من النقط في الإطارين الأيسر العلوي و الأيمن السفلي و بالتالي الحصول على أكبر عدد من النقاط السالبة من الوحدات المعيارية 10 من 13 في المثال السابق. و منه فمن الطبيعي الحصول على معامل ارتباط خطي سالب باعتباره المتوسط الحسابي لهذه القيم.

معامل (التحديد)

من اجل معرفة إذا كانت العلاقة بين متغيرين x و y علاقة خطية ذات دالة إحصائية نحسب قيمة معامل التحديد الذي يساوي مربع معامل الارتباط و يرمز له بالرمز $\mathbf{r^2}$.

في المثال السابق $\mathbf{r = -0.78}$ و منه $\mathbf{r^2 = (-0.78)^2}$

إذن: $\mathbf{r^2 = 0.62}$

الدلالة الإحصائية لمعامل التحديد

لمعرفة ما إذا كانت العلاقة x_i و y_i ذات دلالة إحصائية و أن العلاقة الملاحظة من خلال معامل الارتباط لم تكن نتيجة عوامل الصدفة نتبع الخطوات التالية:

نبحث في الجدول الخاص بقيم معامل التحديد النظرية عن القيمة المقابلة لعدد الوحدات الإحصائية المقاسة

إذا كانت قيمة معامل التحديد المحسوبة أكبر من قيمة معامل التحديد النظرية نقول أن العلاقة بين x_i وy_i ذات دلالة إحصائية.

و إذا كانت قيمة معامل التحديد المحسوبة أقل من قيمة معامل التحديد النظرية نقول أن العلاقة بين x_i وy_i غير دلالة إحصائية و بالتالي المتغيرين x_i وy_i غير مرتبطين.

في المثال السابق نطرح السؤال

هل العلاقة بين السن و نبض القلب الأقصى هي علاقة خطية ذات دلالة إحصائية؟

الإجابة:

معامل التحديد $\mathbf{r^2 = 0.62\ \%}$ و عدد الوحدات الإحصائية المقاسة هو $n = 13$ و بالرجوع الى الجدول نجد أن قيمة $\mathbf{r^2 = 31\%}$ و هي أقل من قيمة معامل التحديد المحسوبة و بالتالي نستطيع القول أن:

العلاقة بين السن و نبض القلب الأقصى هي علاقة خطية ذات دلالة إحصائية.

تمرين:

إليك قيم الوزن و الطول الخاصة بلاعبي المنتخب الجزائري لكرة اليد:

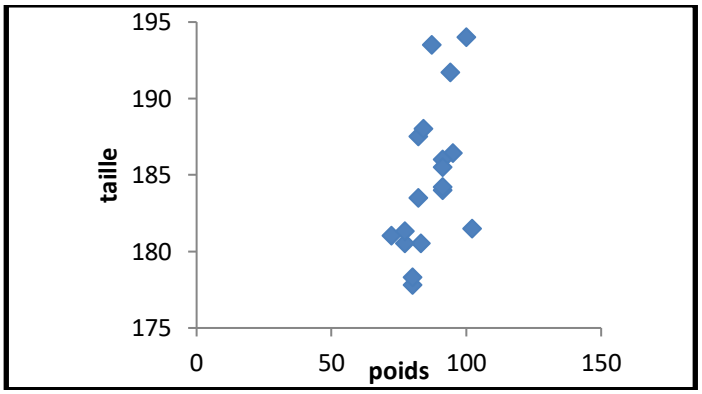
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
الوزن	82	91	102	95	91	83	87	91	84	80	77	82	91	72	94	80	100	77
الطول	187	184	181	186	186	180	193	184	188	177	180	183	185	181	191	172	194	181

- أحسب مؤشر الكتلة الجسمية : $\left(IMC = \frac{poids}{taille^2}\right)$
- أرسم لوحة انتشار قيم الوزن على الطول.
- أرسم لوحة انتشار الطول على IMC و انتشار الوزن على IMC ماذا تستنتج؟
- أدرس العلاقة بين كل من (الوزن على الطول) (IMC والطول) (IMC و الوزن) ماذا تستنتج؟
- أحسب معامل التحديد لمعاملات الارتباط المحسوبة.

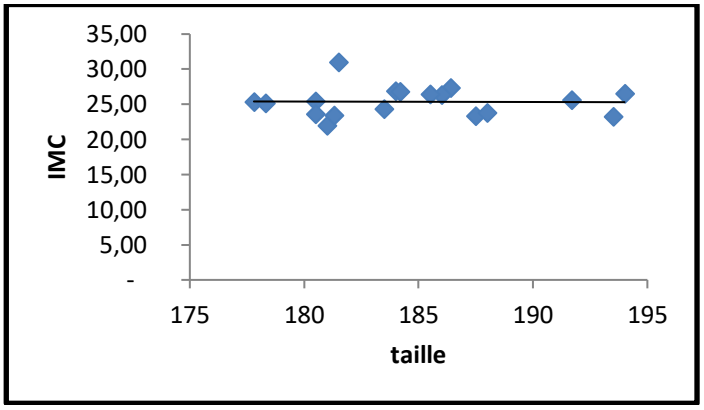
د. يوسف بن شيخ

أستاذ محاضر –أ– بقسم التربية البدنية و الرياضية جامعة باجي مختار عنابة

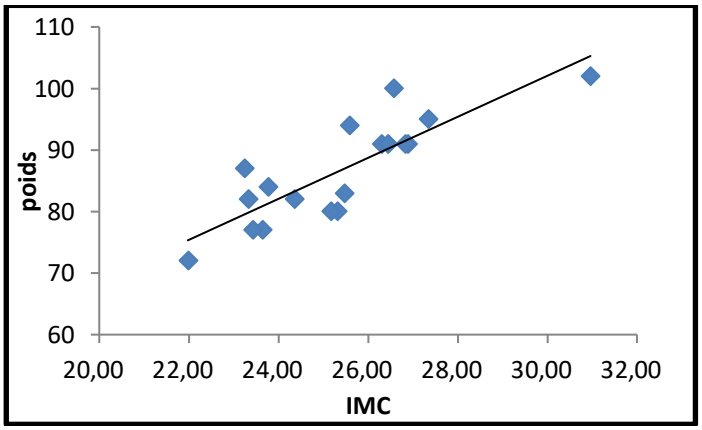
الحل:



1- لوحة انتشار قيم الوزن على الطول



2- لوحة انتشار قيم الطول على IMC



3- لوحة انتشار قيم الوزن على IMC

الاستنتاجات:

من خلال لوحات الانتشار نستنتج ما يلي:

- من خلال لوحة انتشار الوزن على الطول نلاحظ أن السحابة النقطية اتخذت شكل مستقيم عمودي و نلاحظ أن الزيادة في قيمة الوزن لا يتبعها زيادة في الطول
- من خلال لوحة انتشار الطول على IMC نلاحظ أن السحابة النقطية اتخذت شكل مستقيم أفقي و نلاحظ أن الزيادة في قيمة الطول لا يتبعها زيادة في IMC

د. يوسف بن شيخ
أستاذ محاضر –أ– بقسم التربية البدنية و الرياضية جامعة باجي مختار عنابة

- بالنسبة للوحة انتشار قيم الوزن على IMC نلاحظ أن السحابة النقطية اتخذت شكل مستقيم متصاعد و نلاحظ أن الزيادة في قيمة الوزن يتبعها زيادة في قيمة IMC.

إذن مما سبق نستنتج أنه توجد علاقة خطية واحدة كانت بين الوزن و IMC.

دراسة العلاقات:

Poids kg	Taille cm	Taille m	IMC	Z_x	Z_y	Z_y	Z_x*Z_y	Z_x*Z_y	Z_{imc}	$Z_{imc} * Z_y$	$Z_{imc} * Z_x$
82	187,5	1,875	23,32	-0,56	0,5	0,57	-0,32	- 0,28	-0,98	- 0,56	0,54
91	184	1,84	26,88	0,53	-0,2	- 0,15	- 0,08	- 0,11	0,73	- 0,11	0,39
102	181,5	1,815	30,96	1,86	-0,7	-0,67	- 1,24	- 1,30	2,69	- 1,80	5,01
95	186,4	1,864	27,34	1,01	0,28	0,35	0,35	0,28	0,95	0,33	0,97
91	186	1,86	26,30	0,53	0,2	0,26	0,14	0,11	0,45	0,12	0,24
83	180,5	1,805	25,48	- 0,44	-0,9	- 0,87	0,38	0,39	0,06	-0,05	- 0,02
87	193,5	1,935	23,24	0,05	1,7	1,81	0,09	0,08	- 1,02	- 1,85	- 0,05
91	184,2	1,842	26,82	0,53	-0,16	-0,11	- 0,06	- 0,08	0,70	- 0,08	0,37
84	188	1,88	23,77	- 0,32	0,6	0,68	- 0,21	-0,19	-0,77	- 0,52	0,24
80	177,8	1,778	25,31	- 0,80	-1,44	- 1,43	1,14	1,15	- 0,03	0,04	0,02
77	180,5	1,805	23,63	- 1,16	-0,9	-0,87	1,01	1,04	- 0,83	0,73	0,96
82	183,5	1,835	24,35	- 0,56	-0,3	- 0,25	0,14	0,17	- 0,48	0,12	0,27
91	185,5	1,855	26,45	0,53	0,1	0,16	0,08	0,05	0,52	0,08	0,28
72	181	1,81	21,98	- 1,76	-0,8	-0,77	1,36	1,41	- 1,63	1,25	2,87
94	191,7	1,917	25,58	0,89	1,34	1,44	1,29	1,20	0,11	0,15	0,09
80	178,3	1,783	25,16	- 0,80	-1,34	- 1,33	1,06	1,07	- 0,09	0,12	0,08
100	194	1,94	26,57	1,62	1,8	1,92	3,10	2,91	0,58	1,11	0,94
77	181,3	1,813	23,43	- 1,16	-0,74	- 0,71	0,82	0,86	- 0,93	0,66	1,08

من خلال الجدول يتبين ما يلي:

معامل الارتباط بين الوزن و الطول عرف قيمة قدرها **r = 0.50** وهذا يعني وجود علاقة طردية متوسطة بين الطول و الوزن.

معامل الارتباط بين IMC و الطول عرف قيمة قدرها **r = 0.01–** وهذا يعني عدم وجود علاقة بين الطول و IMC

معامل الارتباط بين IMC و الوزن عرف قيمة قدرها **r = 0.79** وهذا ما يدل وجود علاقة طردية قوية بين الوزن و IMC.

حساب معامل التحديد:

معامل التحديد بين الوزن و الطول $r^2 = (0.5)^2$ و منه $r^2 = 0.25$

معامل التحديد بين IMC و الطول $r^2 = (-0.01)^2$ و منه $r^2 = 0$

معامل التحديد بين الوزن و IMC $r^2 = (0.79)^2$ و منه $r^2 = 0.61$

بالرجوع إلى الجدول نلاحظ أنه عند n= 18 $r^2 = 22$

ومنه نستنتج أن العلاقة بين الوزن و IMC علاقة خطية ذات دلالة إحصائية.