Collage of Engineering

Dams and Water Resources Eng. Department

Second Class

Statistical Engineering Lectures

Prepared by: Lecturer.Dr. NAHLA N. HILAL

الفصل الاول

الاحصاء: *يعني جمع البيانات التي تخص تعداد النفوس وعدد الولادات والوفيات او بيانات تخص الطقس كدرجات الحرارة والرطوبه النسبيه....الخ وبعدها اخذ يشمل عرض هذه البيانات بصورة مركزة في جداول ورسوم بيانية ،ولكن علم الاحصاء الحديث مرتبط بالطرق العلمية التي تخص جمع وتنظيم وتلخيص وعرض البيانات من اجل الاستدلال لاتخاذ القرارات .*

يقسم علم الاحصاء الى قسمين:

1- *الاحصاء الوصفي(Descriptive statistics):يشمل جمع وتبويب البيانات.*
2- *الاحصاء الاستدلالي(Inferential statistics):يعني الاستدلال من النتائج لاتخاذ القرارات ويعتمد بدرجة كبيرة على نظرية الاحتمال ويتممها.*

تطبيقات علم الاحصاء:ان لعلم الاحصاء تطبيقات شتى في كافة المجالات ،بضمنها المجال الهندسي وعلى سبيل المثال في مجال الهندسة المدنية يستخدم في التصاميم الانشائية ودراسة طبيعة حركة المرور لايجاد السبل المناسبة للسيطرة عليهاوفي تصميم بوابات سدود الخزن .

ان لتفسير اي احتمال هناك موقفين :

1. *الموقف الموضوعي (Objective position)*
2. *الموقف الشخصي (Personalistic position)*

يطبق الموقف الموضوعي على الحوادث التي تكرر نفسها مثل انتاجية القوالب الخرسانية في معمل الابنية الجاهزة.

اما الموقف الشخصي فانه يعتمد جزئيا على الرأي الشخصي، وفي منهاجنا لانتعامل مع هذا الموقف.

المصطلحات الاحصائية

العينة (Sample): *العينة تمثل مجموعة معينة من عناصر المجتمع .فالعشرون قالبا من الجدران الخرسانية تعتبر عينة لانتاجية المعمل لمدة سنة او اكثر والمئة فرد في المدينة يعتبرون عينه للسكان .*

العينة العشوائية (Random Sample) : *تعتبر العينة عشوائية اذا اعطي كل عنصر من مكونات المجتمع الفرصة نفسها في الاختيارليكون جزء من هذه العينة,وقد يكون هذا العنصر قوة تحمل ناتجة عن فحص قالب خرساني من انتاجية معمل الابنية الجاهزة.*

العينات المستقلة (Independent Sample) : *يقال للعينتان انهما مستقلتان اذا تم اختيارهما بصورة عشوائية اي تكون كل واحدة منهما عينة عشوائية.*

المتغير الاحصائي والمتغير المحدد (Statistical Variable and Variate): *المتغير الاحصائي يشير الى صفة معينة للبيانات ويرمز له ب (X)اما المتغير المحدد فانه القيمة الخاصة من المتغير ويرمز(x) فقد يكون ممثلا للاجهاد ا والانفعالاو التصريف في الانهار ويكون (X) ممثلا للقيمة التي ياخذها هذا المتغير.*

المتغير العشوائي (Random Variable) : *قيمة المتغير العشوائي رقماً حصل نتيجة تجربة ويكون متقطعا او متصلاً*

فالمتغير المتقطع يكون لعينته مجالا منتهيا ويسمى المتغير العشوائي المنفصل او المميز فمثلاً عدد الطلبة في قاعة المحاضرات او عدد المركبات على جسر معين تعتبر متغيرات متقطعة , لان المتغير المحدد المحدد ياخذ مجالا محدداً او ارقاما محددة على العكس من مجال عينة المتغيرالعشوائي المتصل فانه غير متصل وياخذ اي قيمة ضمن مدى تراوح القيم .

المتغير المستقل او التابع(Independent and Dependent Variable): *يقال للمتغيرين مستقلان اذا لم تتأثر نتيجة احدهما بتثبيت قيمة الآخر .فعند رمي قطعتي النقود في الفضاء فان الاتجاه الذي تاخذه القطع الاولى عند استقرارها على الارض لايؤثر على الاتجاه الذي تاخذه قطعة النقود الثانية .ان الاتجاه هو المتغير والقطعتان هما المتغيران المستقلين .ولكن عند دراسة الاجهاد والانفعال في عمود خرساني فان لكل قيمة من الاجهاد ,ياخذ الانفعال نتيجة او قيمة معينة , فيكون الاجهاد متغيرا مستقلا ,والانفعال متعيرا تابعا .*

الفصل الثاني

التوزيعات

يتضمن هذا الفصل طرق عرض البيانات الخام بصورة مركزة لتوضيح توزيع وانتشار هذه البيانات.ولغرض اعطاء صورة واضحة للبيانات الخام فقد تمثل برسم بياني او يتم تنظيمها بجداول فالرسم البياني ماهو الا عرض صوري لتوضيح العلاقة بين المتغيرات ،ويكون اما خارطة او مخطط . الجدول(2-1) ادناه يمثل سرع مركبات مقاسه الى اقرب 1كم/س لعينه على طريق المثنى بغداد وهذه طريقه مبسطه لعرض البيانات بشكل صفوف مرتبه تصاعديا او تنازلياً .

الجدول (2-1): بيانات السرع بصورة عشوائية

37	61	76	40	54	74	37	48	47	53
40	63	63	68	57	55	59	54	52	56
87	74	51	54	57	59	46	41	44	58
65	67	64	60	82	51	50	54	51	55
67	57	59	84	66	50	50	56	46	***32***
47	45	61	40	63	60	53	54	52	51
70	45	73	76	67	43	50	61	71	55
57	53	65	61	55	41	77	56	64	52
36	50	59	62	42	72	73	68	48	69
46	55	60	70	70	58	65	53	71	78

كما مبين من الجدول اعلاه بان اقل قيمة هي(32) واعلى قيمة هي (87) فالفرق بين هاتين القيمتين يمثل المدى (rang) فالمدى يكون 87-32=55.

الجدول (2-2) : بيانات السرع بصورة تصاعدية

32	36	37	40	40	40	41	41	42	43
44	45	45	46	46	46	47	47	48	48
50	50	50	50	50	51	51	51	51	52
52	53	53	53	53	54	54	54	54	54
55	55	55	55	55	56	56	56	57	57
57	457	58	58	59	59	59	59	59	60
60	60	61	61	61	61	62	62	63	63
63	64	64	65	65	65	66	67	67	67
68	68	69	70	70	70	71	71	72	73
73	74	74	76	76	77	78	82	84	87

الجدول (2-3) : جدول التوزيع التكراري للبيانات

الصنف	التكرار علامة	التكرار رقما
31-35	I	1
36-40	I I I I I	5
41-45	I I I I I I I	7
46-50	I I I I I I I I I I I I	12
51-55	I I I I I I I I I I I I I I I I I I I I	20
56-60	I I I I I I I I I I I I I I I I I	17
61-65	I I I I I I I I I I I I I I	14
66-70	I I I I I I I I I I	10
71-75	I I I I I I I	7
76-80	I I I I	4
81-85	I I	2
86-90	I	1

من الجدول(2-3) يتبين ان مجموع التكرارات يساوي مجموع القراءات ويساوي 100.
طول الصنف هو الفرق بين النهايتين الحقيقيتين للصنف .اما مركز الصنف فهو معدل نهايتي الصنف .

الجدول(2-4)

Class i	Class limit	True limit	Class Center-xi	Frequency-fi	Probability Pi=fi/n
1	31-35	30.5-35.5	33	1	0.01
2	36-40	35.5-40.5	38	5	0.05
3	41-45	40.5-45.5	43	7	0.07
4	46-50	45.5-50.5	48	12	0.12
5	51-55	50.5-55.5	53	20	0.20
6	56-60	55.5-60.5	58	17	0.17
7	61-65	61-65.5	63	14	0.14
8	66-70	66-70.5	68	10	0.10
9	71-75	71-75.5	73	7	0.07
10	76-80	75.5-80.5	78	4	0.04
1	81-85	80.5-85.5	83	2	0.02
12	86-90	85.5-90.5	88	1	0.01

من الجدول(2-4).

- طول الصنف هو الفرق بين النهايتين الحقيقيتين للصنف و 5 =40.5-35.5.
- اما مركز الصنف فهو معدل نهايتي الصنف.
- ومجموع الاحتمالية =1.

***التوزيعات الاحصائية*:** *يمثل جدول التكرار المكون من الاصناف،الممثلة بنهاياتها او مراكزهااو نهايتهاالحقيقية،وتكرار تلك الاصناف،توزيعا احصائيا،حيث يبين الجدولنمط توزيع البيانات بصوره واضحة ومن اهم التوزيعات الاحصائية المستخدمة لهذا الغرض والمسخرجة من جدول التكرار:*

1.التوزيع التكراري: الجدول الموضح لنمط تغيير التكرار مع الاصناف الممثلة بمجالاتها،نهاياتها الحقيقيه او مراكزهاويرمز لدالة هذا التوزيع ب (f_x).

2.التوزيع الاحتمالي :الجدول الموضح لنمط تغيير الاصناف ويرمز لدالته ب (p_x).

3. التوزيع التكراري المتراكم : هو الجدول الموضح لعدد القراءات الاقل من قيمة معينة ويفضل ان تكون الاخيرة النهايات الحقيقية للاصناف ويسمى التوزيع المتراكم الصاعد او توزيع اقل من ، اما الجدول الموضح لعدد القراءات المساوية الاكبر من قيم معينه،فيطلق عليه التوزيع المتراكم النازل او توزيع اكبر من. وكما موضح بالجدول رقم**(2-5)**

الصنف	مركز الصنف	التكرار	الاحتمال pi	الصنف	التكرار المتراكم الصاعد	المتراكم (p_x)	الصنف	التكرار المتراكم النازل
30.5-35.5	33	1	0.01	اقل من 35.5	0	0.01	اكثر من 30.5	100 99
35.5-40.5	38	5	0.05	اقل من 40.5	1	0.06	اكثر من 35.5	94
40.5-45.5	43	7	0.07	اقل من 45.5	6	0.13	اكثر من 40.5	87
45.5-50.5	48	12	0.12	اقل من 50.5	13	0.25	اكثر من 45.5	75
50.5-55.5	53	20	0.20	اقل من 55.5	25	0.45	اكثر من 50.5	55
55.5-60.5	58	17	0.17	اقل من 60.5	45	0.62	اكثر من 55.5	38
61-65.5	63	14	0.14	اقل من 65.5	62	0.76	اكثر من 61	24

66-70.5	68	10	0.10	اقل من 70.5	76	0.860	اكثر من 66	14
71-75.5	73	7	0.07	اقل من 75.5	86	0.93	اكثر من 71	7
75.5-80.5	78	4	0.04	اقل من 80.5	93	0.97	اكثر من 75.5	3
80.5-85.5	83	2	0.02	اقل من 85.5	97	0.99	اكثر من 80.5	1
85.5-90.5	88	1	0.01	اقل من 90.5	100	1.00	اكثر من 85.5	0

مثال:اجريت دراسة لمعرفة مدى استغلال المركبات المستخدمة لموقف الجامعة وتم اخذ عينة بحجم 50مركبة في صباح احد الايام ،وكان عدد الاشخاص في كل مركبه كالاتي :

1 1 2 1 4 2 2 5 1 3

3 1 1 2 2 1 3 1 2 2

1 4 3 3 1 4 2 2 1 3

3 2 1 2 3 2 1 3 2 4

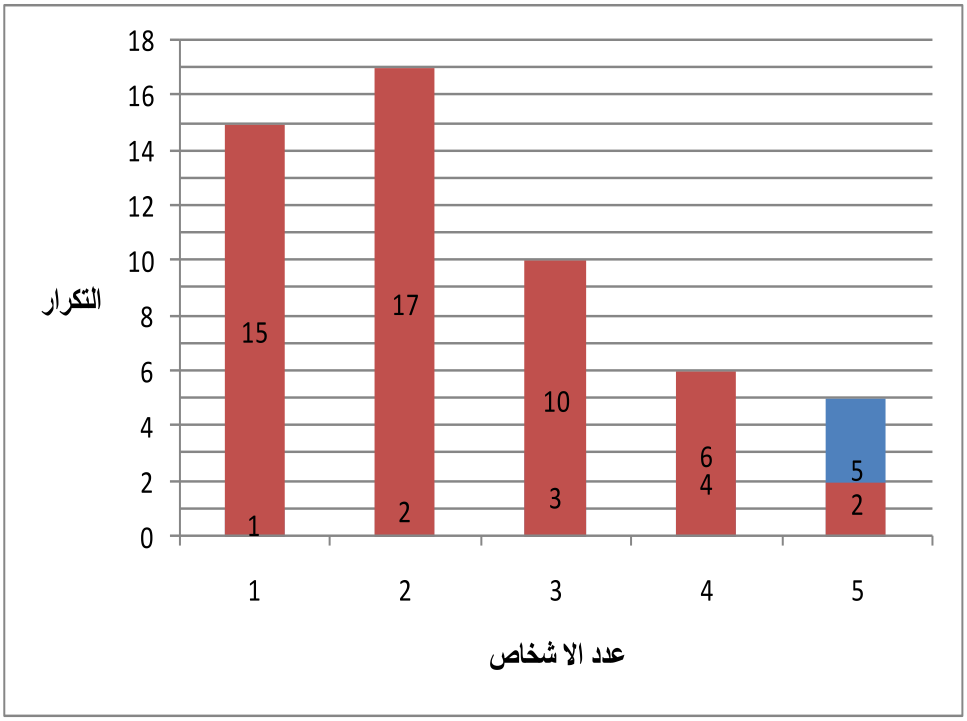
4 5 1 1 2 3 2 4 2 2

كون جدول التكرار لهذه البيانات وارسم المدرج؟

1- جدول التكرار

الصنف عدد الاشخاص	التكرار
1	15
2	17
3	10
4	6
5	2

رسم المدرج


18
16
14
12
10
8
6
4
2
0
التكرار
15
17
10
6
4
5
1
2
3
2
1
2
3
4
5
عدد الا شخاص

المحاضرة الثانية

مثال: تم قياس السقوط المطري السنوي في الصحراء الغربية خلال فترة سنة وكانت القراءات مقاسة الى اقرب 1 ملم كالاتي :

11 13 16 17 19 20 22 22 23 25
26 26 27 28 30 31 32 36 37 42

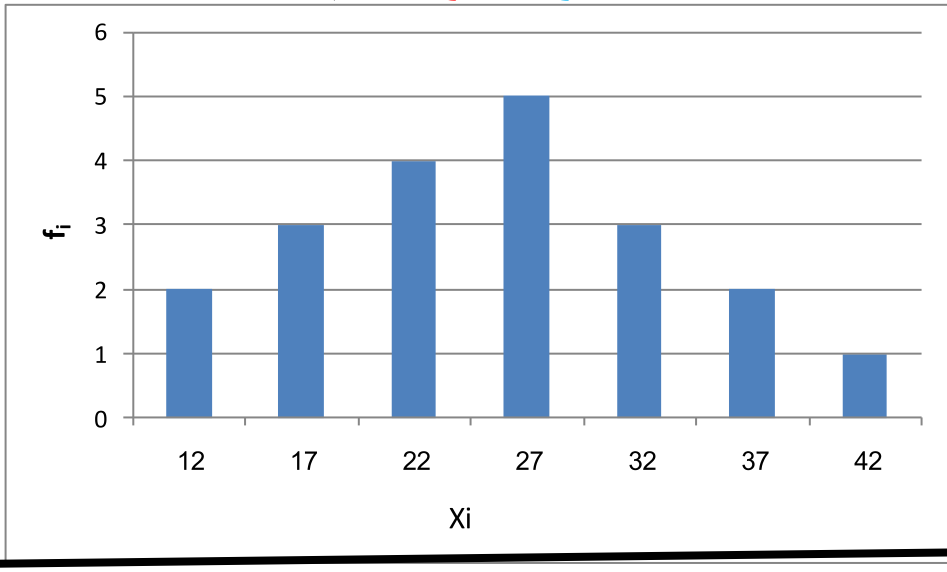
1-كون جدول التكرار مبتدأ بالصنف 10-14.
2 - جد مدرج ومضلع التكرار والاحتمال وتوزيع الاحتمال المتراكم الصاعد.

3-احتمال سقوط مطري سنوي اقل من 24.5 ملم .
4-احتمال سقوط مطري بين 24.5و 34.5 ملم.
5- احتمال سقوط مطري 34 ملم.

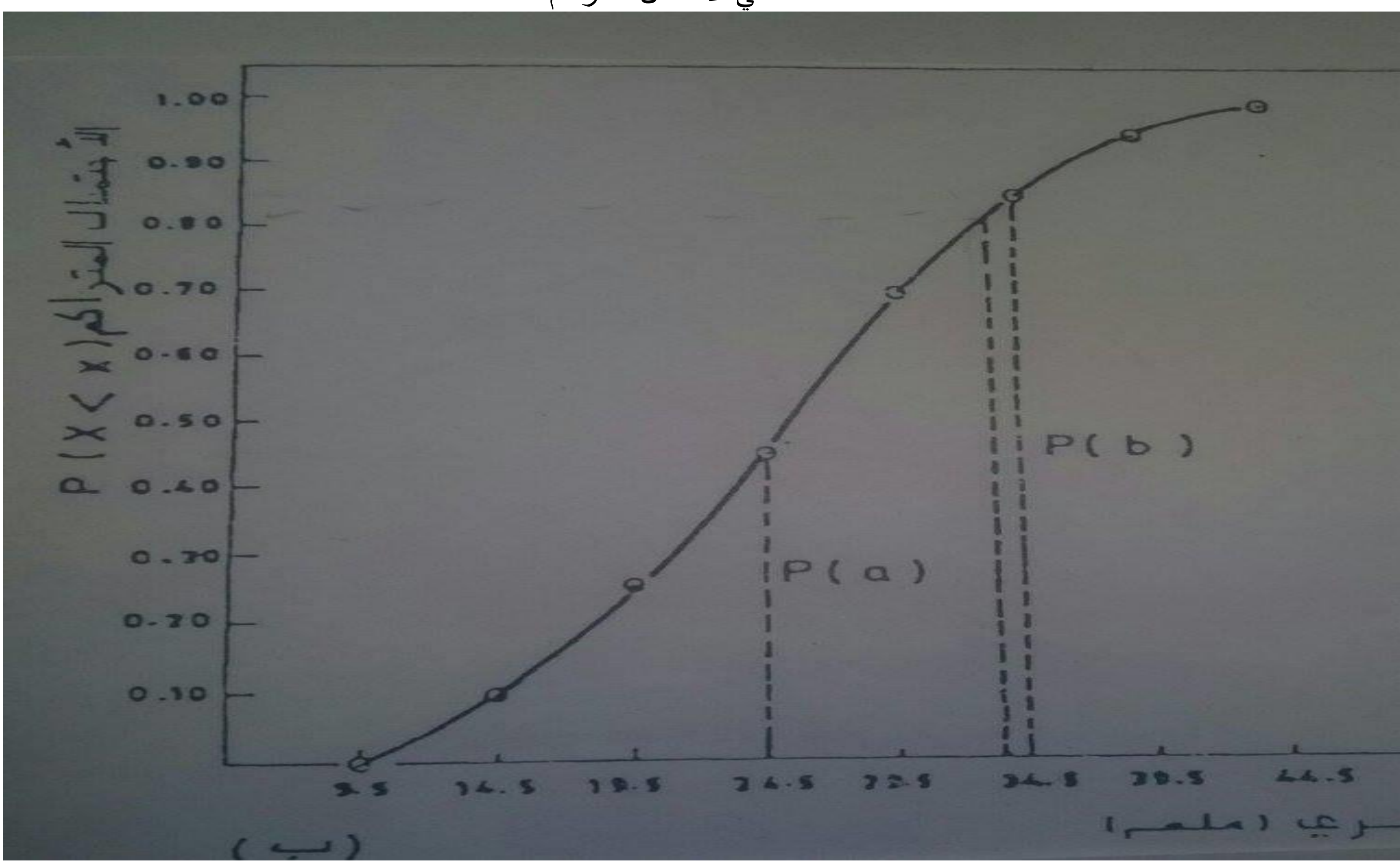
Class	Xi	f_i	P_i	Class	F(x)	P(x)
				9.5 اقل من	0	0
10-14	12	2	0.10	14.5	2	0.1
15-19	17	3	0.15	19.5	5	0.25
20-24	22	4	0.20	2.5	9	0.45
25-29	27	5	0.25	29.5	14	0.7
30-34	32	3	0.15	34.5	17	0.85
35-39	37	2	0.1	39.5	19	0.95
40-44	42	1	0.05	44.5	20	1.00
		$\sum= 20$	$\sum=1$			

- المدى الفرق بين اعلى واقل قيمة في البيانات.
- عدد الفئات ناخذه بين7-15 حسب الخبرة
- طول الفئة نجده من قسمة المدى على عدد الفئات ويفضل ان يكون عدد صحيح.

المدرج والمضلع التكراري



منحني الاحتمال المتراكم



من منحني الاحتمال المتراكم احتمال سقوط مطري اقل من 24.5 ملم

3-$P(X<24.5) = 0.45$

4-$P(24.5<X<34.5) = 0.85-0.45= 0.4$

5-$P(X<34.5)- P(X<33.5)= 0.85-0.82=0.03$

H.W: 1

الجدول التالي يوضح نتائج الفحص لمقاومة الانضغاط ل 45 مكعب خرساني مقاس الى اقرب عشر نت/ملم^2

1-كون جدول التكرار.

2-ارسم مدرج ومضلع الاحتمال والتكرار.

3-جد النسبة المئوية للمكعبات بقوة انضغاط اقل من 41.0 نت/ملم^2 .

4- جد النسبة المئوية للمكعبات بقوة انضغاط اقل من 28.5 نت/ملم^2 او اكبر.

24	29	32	25.4	24.8	34.2	33.8	37.8	44.4
28	28	29	35.0	24	29	23	23	23
37	25	291	36.4	29	22	22	24.2	35.2
34	15	24.4	27.4	31	38.6	28.5	22	29
33.8	19.2	26	24.4	19	34.8	33	14	33.2

المحاضرة الثالثة
مقاييس موقع التمركز

وهي مقاييس تستخدم لايجاد القيمة النموذجية الممثلة للبيانات والتي يشار اليها بقيمة التمركز لذا تسمى المتوسطات بمقاييس النزعه المركزية وهناك انواع عديدة لهذه المتوسطات منها :-

1-الوسط.
2-الوسيط .
3-المنوال .
4-منتصف المدى .

الوسط :وهو اكثر القيم استخداما للتعبير عن موقع التمركزللبيانات حيث توجداربع طرق للتعبير عنه.

- ***الوسط الحسابي.***
- ***الوسط الهندسي.***
- ***الوسط التوافقي .***
- ***جذور وسط المربعات .***

الوسط الحسابي للبيانات الغير مبوبة ($\bar{X}$):

$$\bar{X} = \sum_{i=1} X_i / n$$

مثال: اذا كانت درجات احد الطلاب في مادة حقوق الانسان لثلاثة امتحانات كالاتي فجد الوسط الحسابي ؟ 80,75,95.

$\sum X_i = 80+75+95 = 250$

$n=3$
$\bar{X} = \sum X_i / n = 250/3 = 83.33.$

الوسط الحسابي للبيانات مبوبة ($\bar{X}$) :

$$\bar{X} = \sum X_i f_i / \sum f_i$$

الوسط الهندسي للبيانات الغير مبوبة ($\log \bar{x}$) :

$$\log \bar{X}_g = 1/n \sum_{i=1}^{n} \log x_i$$

الوسط الهندسي للبيانات المبوبة ($\log \bar{x}$)

$$\log \bar{X}_g = 1/n \sum_{i=1}^{n} f_i \log x_i$$

مثال: اذا كانت نفاذية التربة باتجاه المحاور X,Y,Z [$2.7*10^{-5}$ m/sec] [$2.0*10^{-5}$ m/sec] [$5.0*10^{-5}$ m/sec]

على الترتيب جد الوسط الهندسي؟

$$\text{Log } \bar{X}_g = 1/n \sum_{i=1}^{n} \log x_i$$

$$\bar{X}_g = [2.7*10^{-5} * 2.0*10^{-5} * 5.0*10^{-5}]^{1/3} = 3.0*10^{-5} \text{ m/sec.}$$

الوسط التوافقي للبيانات الغير مبوبه(X_h) :

$$X_h = n / \sum_{i=1}^{n} 1/x_i$$

الوسط التوافقي للبيانات مبوبه(X_h) :

$$X_h = \sum_{i=1}^{n} f_i / \sum_{i=1}^{n} f_i/x_i$$

مثال :جد الوسط الحسابي والهندسي والتوافقي للبيانات المدونة بالجدول ادناه؟

Class	fi
10-14	2
15-19	3
20-24	4
25-29	3
30-34	2
35-39	1

لغرض الحل نرتب الجدول التالي:

Class	xi	fi	xi fi	log xi	fi log xi	fi / xi
10-14	12	2	24	1.079	2.158	0.166
15-19	17	3	51	1.230	3.691	0.176
20-24	22	4	88	1.342	5.369	0.181
25-29	27	3	81	1.431	4.294	0.111
30-34	32	2	64	1.505	3.01	0.062
35-39	37	1	37	1.568	1.568	0.027
∑	147	15	345		20.090	0.725

1- $\bar{X}=\sum X_i f_i / \sum f_i=345/15=23.$

2- $\text{Log } \bar{X}_g=1/n \sum_{i=1}^{n} f_i \log x_i =21.88.$

3- $X_h=\sum_{i=1}^{n} f_i / \sum_{i=1}^{n} f_i/x_i=15/0.725.$

الوسيط: يتم ترتيب البيانات تصاعديا او تنازليا فاذا كان عدد البيانات زوجي ناخذ معدل الرقمين الوسطيين ليمثل الوسيط اما اذا كان عدد البيانات فرديا يؤخذ الرقم الوسطي ليمثل الوسيط.

مثال :

بما ان عدد البيانات فردي فالوسيط هو 2 1.0,1.5,2.0,2.5,3.0

بما ان عدد البيانات زوجي فالوسيط هو 1.0,1.5,2.0,2.5,3.0,3.5

2+2.5/2= 2.25

الوسيط للبيانات المبوبة

الحل : ان قيمة الوسيط للبيانات يساوي وسط القيمتين العاشره والحادية عشر لكون عدد البيانات زوجيا

Class	fi
10-14	2
15-19	3
20-24	4
25-29	3
30-34	2
35-39	1

M=25+26/2=25.5

$M=a+ (n/2-n1)/f_m. \Delta$
M=24.5+ (10-9/5)/5
M=25.5

Δ و a=يمثلان طوله ونهايته الحقيقية والدنيا .
n=عدد التكرارات.
n1= التكرار المتراكم للصنف الذي يسبق الصنف الوسطي مباشرة.
f_m= تكرار الصنف الوسطي.

المنوال :هو القيمة الاكثر تكرارا للبيانات الغير المبوبة .

1, 1.4,2,3,5,6,6,4,4

المنوال = 4.0

المنوال :هو القيمةالتي تمتلك اكبر تكرارا للبيانات المبوبة .

Class	fi
2	2
4	1
6	4
8	2
11	1
12	5
16	3
18	1

لمنوال = 12.0

HW: 2

24	29	32	25.4	24.8	34.2	33.8	37.8	44.4
28	28	29	35.0	24	29	23	23	23
37	25	291	36.4	29	22	22	24.2	35.2
34	15	24.4	27.4	31	38.6	28.5	22	29
33.8	19.2	26	24.4	19	34.8	33	14	33.2

للجدول اعلاه

1-كون جدول التكرار.
2- جد الوسط الحسابي والهندسي والتوافقي.
3- جد الوسيط والمنوال.

المحاضرة الرابعة

مقاييس التغير: *هي المقاييس التي تبين مدى تغير البيانات فيما بينها ومدى التفاوت والتغيير في المجموعة المعنية من هذه البيانات.*

- *هنا ك تعابير مختلفة تستخدم لوصف درجة تغيير البيانات حول الوسط اهمها:*
  1. *المدى.*
  2. *التقسيمات الجزيئية.*
  3. *وسط الانحراف.*
  4. *التباين.*
  5. *معامل التغير.*
  6. *معامل التغير الربيعي.*

المدى (The Range) :هو الفرق بين اعلى واقل قيمة في البيانات'فبالرغم من سهولة استخراجه الا ان المدى لايعطي صورة واضحة عن التغير في عموم البيانات لانه يعتمد على القيمتين المتطرفتين فقط ،ولايعتمد على جميع البيانات.

80 85 90 98 104 115 122 130 (X=103)

80 99 100 102 103 104 106 130(X=103)

- نلاحظ ان المدى للمجموعه الاولى =المدى للمجموعه الثانية =50 ،الاان تغيير بيانات المجموعه الاولى حول الوسط اعلى من تغير بيانات المجموعه الثانية.

مدى التقسيمات الجزيئية (Quantiles Range) : يمكن استخدام مدى حدود التقسيمات الجزيئية لحصر البيانات المتطرفة الصغيرة والكبيرة وابرز هذه التقسيمات المستخدمة هي الربيع الاول و الثالث حيث يكتب المدى (Q_3-Q_1).
للمثال المعطى في المحاضرة الثالثة جد المدى الربيعي والمئيني.

يمكن استخدام منحنى الاحتمال المتراكم (الصاعد)

P_{90}=36.5 P_{10}=14.5 , Q_1=19.5 Q_3=31.

Q_3- Q_1=19.5-14.5=5.0 المدى الربيعي

P_{90} - P_{10}=36.5- 14.5=22.0 المدى المئيني

وسط الانحراف ***(Mean Deviation)***: يستخدم لوصف التغيير حول الوسط ويعرف وسط الانحراف بالمعادلة التالية.

$$m.d = \sum |X_i - \bar{X}| / n \longleftarrow$$ للبيانات الغير مبوبة

$$m.d = \sum |X_i - \bar{X}| f_i / \sum f_i \longleftarrow$$ للبيانات المبوبة

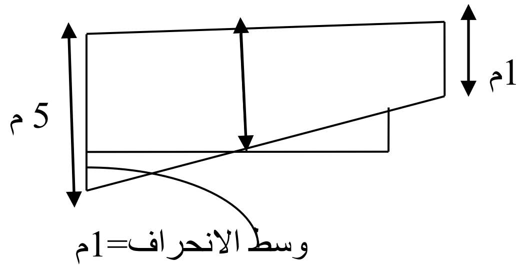
احسب وسط الانحراف للبيانات المعطاة في الجدول التالي:

CLASSES	X_i	f_i	$X_i f_i$	$X_i - \bar{X}$	$\lvert X_i - \bar{X} \rvert$	$\lvert X_i - \bar{X} \rvert f_i$
10-14	12	2	24	-13.5	13.5	27.0
	17	3	51	-8.5	8.5	25.5
	22	4	88	-3.5	3.5	14.0
	27	5	135	1.5	1.5	7.5
	32	3	96	6.5	6.5	19.5
	37	2	74	11.5	11.5	23.0
	42	1	42	16.5	16.5	16.5
∑		**20**	**510**			**133**

$$\bar{X} = \sum X_i f_i / \sum f_i = 510/20 = 25.5$$

$$m.d = \sum |X_i - \bar{X}| f_i / \sum f_i = 133/20 = 6.65$$

☒ للتغير الخطي يكون وسط الانحراف ربع الفرق بين اعلى وادنى قيمة فمثلا عندما يتغيير عمق التربة خطيا من 1-5 متر يكون وسط الانحراف. م 1 =4/ 1-5



التباين (The Variance): كوسيلة للتحلص من الاشارة السالبة ،تربع الانحرافات عن قيمة الوسط وتجمع ثم يؤخذ وسطها للتبير عن التباين ،لابراز درجة تغيير هذه البيانات حيث يرمز له ب (S^2) بالنسبة لبيانات العينه وبالحرف اللاتيني(σ^2) بالنسبة لبيانات المجتمع ويعرف التباين رياضياً للعينة بالمعادلة :

$$S^2= (1/n) \sum_{i=1}^{n} (X_i-\bar{X})^2$$

n =حجم العينه .

$\bar{X}$ = الوسط الحسابي للعينة .

$$S^2= (1/N) \sum_{i=1}^{N} (X_i-\mu)^2$$

ويعرف التباين رياضياً للمجتمع بالمعادلة:

N=حجم المجتمع

µ=الوسط الحسابي للمجتمع

مثال:احسب التباين والانحراف المعياري للبيانات 3.0, 3.5,4, 4.5, 5.0.

Xi	$(X_i-\bar{X})$	$(X_i-\bar{X})^2$
3.0	-1.0	1.0
3.5	-.05	0.25
4.0	0.0	0
4.5	0.5	0.25
5.0	1.0	1
∑ **20**	**0.0**	**2.5**

$S^2= (1/n) \sum (X_i-\bar{X})^2=1/5*2.5=0.5$

$m.d=1/\sqrt{2}$

$$S^2= (1/n-1) \sum_{i=1}^{n} (X_i-\bar{X})^2 \longrightarrow$$ *الانحراف المعياري الغير منحاز* *n ‹ 30*

$$S= [(\sum_{i=1}^{k} (X_i-\bar{X})\, f_i)^2/ \sum_{i=1}^{k} f_i]^{1/2} \longrightarrow$$ معادلة ايجاد الانحراف المعياري للبيانات المبوبة

معامل التغير(Coefficient Of Variation): غالبا مايستخدم المقياس النسبي للتغير بدلا من التغير المطلق لفقرات لابراز اهمية التغيرنسبة الى الاحصائيات الاخرى ولضرورة الحصول على مقياس لابعدى.

$C_V = S/\bar{X}$ للعينة

$C_V = \sigma/\mu$ للمجتمع

معامل التغير الربيعي(***Coefficient Of Quartile Variation***): يعرف معامل التغير الربيعي بالمعادلة

$$C_{q.v} = (Q_3 - Q_1)/(Q_3 + Q_1)$$

مثال: اذا كانت المجموعه الاولى من البيانات $Q_3 = 35, Q_1 = 20$

وللمجموعة الثانية من البيانات $Q_3 = 8, Q_1 = 3$ فبين ايهما اكثر تغيرا؟

$C_{q.v} = 35-20/35+20 = 15/55 = 3/11$

$C_{q.v} = 8-3/8+3 = 5/11$

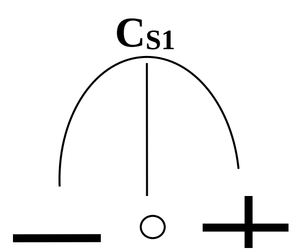
☒ المجموعة الثانية هي ي المجموعه الاكثر تغيرا.

مقاييس الالتواء (Measures Of Skewers): هناك مقياسان رئيسيان للالتواء ،تبرز درجة تناظر التوزيع ،وتبين جهة الالتواء الى اليمين او نحو اليسار التي تعرف بالمعادلة: ويمثل قيمة الالتواء.

للعينة $a = (1/n)\sum (X_i - \bar{X})^3$

للمجتمع $a = (1/N)\sum (X_i - \mu)^3$

اما المقياس الثاني فهو معامل الالتواء ويعرف بالمعادلة: $C_{S1} = a/S^3$



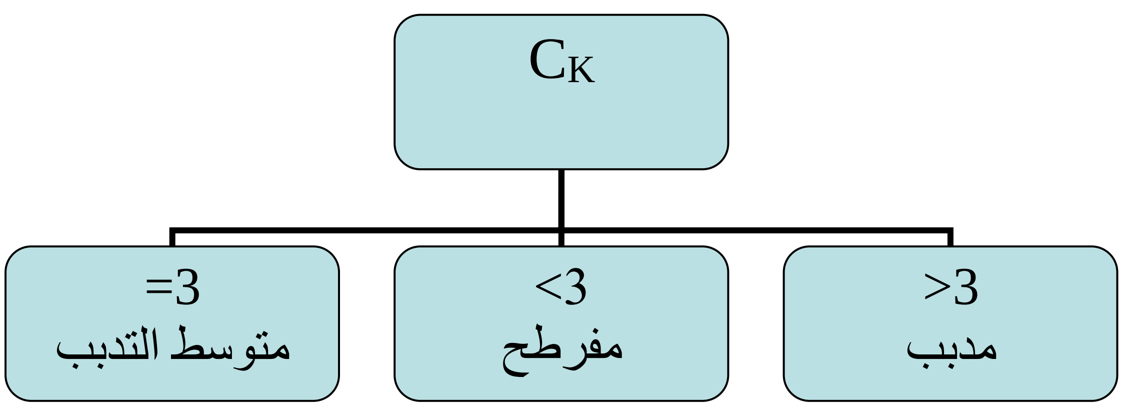
0= التوزيع متناظر

=قيمة سالبة التوزيع ملتو نحو اليسار

=قيمة موجبة التوزيع ملتو نحو اليمين

مقياس التفرطح(Measure of Kertosis): يعرف التفرطح بدرجة تدبب التوزيع وبصورة عامة يؤخذ نسبة التوزيع الطبيعي ويعبر عنه بمعامل التفرطح (Coefficient of *Kertosis*) .

$C_K = [\sum (X_i - \bar{X})^4/n]/S^4$



مثال:اختبار الكثافة النوعية لعشرين عينه من الرمل في مختبر التربة اعطى النتائج التالية في الجدول ادناه.....جد : الوسط الحسابي،التباين والانحراف المعياري الغير منحازين ،معاملات التغير والتفرطح والالتواء.

الصنف	التكرار
2.30-2.39	1
2.40-2.49	2
2.50-2.59	3
2.60-2.69	6
2.70-2.79	7
2.80-2.89	1

x_i	f_i	$x_i f_i$	$(X_i-\bar{X})$	$(X_i-\bar{X})^2 f_i$	$(X_i-\bar{X})^3 f_i$	$(X_i-\bar{X})^4 f_i$
2.345	1	2.345	-0.295	0.08702	-0.02572	0.00757
2.445	2	4.890	-0.195	0.07605	-0.01483	0.00289
2.545	3	7.635	-0.095	0.02708	-0.00257	0.00024
2.645	6	15.870	+0.005	0.00015	0.000001	0.00000
2.745	7	19.215	+0.105	0.07718	0.00810	0.00085
2.845	1	2.845	+0.205	0.04202	0.00862	0.00177

$\bar{X} = \sum X_i f_i / \sum f_i = 2.64$

$S^2 = (1/n) \sum (X_i-\bar{X})^2 = 0.0163$

$S = \sqrt{0.0163} = 0.1276$

$C_V = S/X = 0.0483$

$C_S = a/S^3 = -0.6354$ التوزيع ملتو لليسار

$C_K = [\sum (X_i-\bar{X})^4/n] / S^4 = 2.5123$ التوزيع مفرطح

المحاضرة الخامسة
مبادئ نظرية الاحتمال

احتمال وقوع الحادثة يسمى بالنجاح وعدم حدوثها يسمى بالفشل فعند رمي زهر النرد يمكن ظهور واحد من ست نتائج ،لذلك يمكن احتمال الحصول على اية منهن في الرمية الواحدة **(1/6)** . واذا حدثت حادثة كل عشر سنوات خلال سنوات سابقة فان احتمال حدوثها في اية سنة لاحقه يكون **(1/10)** وتسمى الفترة **(T)** التي تحدث فيها الحادثة بفترة الرجوع ويكون الاحتمال مساويا الى **(1/ T)**. من هذا نلاحظ ان احتمال وقوع المتغير المحدد**(X)** في اي صنف من الاصناف ،هو نسبة تكرار الصنف الى التكرار الكلي ، واشير له ب **P (x)** وعلية فان الاحتمال يساوي التكرار النسبي وان الاحتمال التراكمي يساوي التكرار التراكمي النسبي .لذا يكون احتمال وقوع الحادثة **E مساويا الى نسبة التكرار بالحادثة الى التكرار الكلي .**

<u>نظرية الاحتمال:</u>

☒ اذا كان حاصل تجربة العدد n من النتائج و m منها مصاحب للحادثة E فيكون احتمال وقوع الحادثة : **P(E)= m/n**

واذا كانت E^{-} تشير الى عدم حدوث الحادثة E فان : **$P(E^{-})=(n-m)/n=1-m/n=1-P=q$**

☒ **لذا فأن :**

$$P(E)+P(E^{-})=P+q=1.0$$

☒ **يتبين من هذا ان احتمال وقوع الحادثة يكون رقما بين 0.0 و1.0.**

<u>مثال :</u> لاستغلال المياه الجوفية في احد السهول ،تم حفر 150بئرا فكان 75 بئرا عميق و 90 بئرا عذبا و50 بئرا ارتوازيا.عند القيام بحفر بئر جديد ،ماهو الترجيح لصالح بئر عذب ؟

<u>الحل</u>

- احتمال الحصول على بئر عذب

P (E) = m/n=90/150=3/5.

- احتمال الحصول على بئرغير عذب

$P(E^{-})=(n-m)/n=150-90/150=2/3.$

- وعلية فان الترجيح لصالح بئر عذب

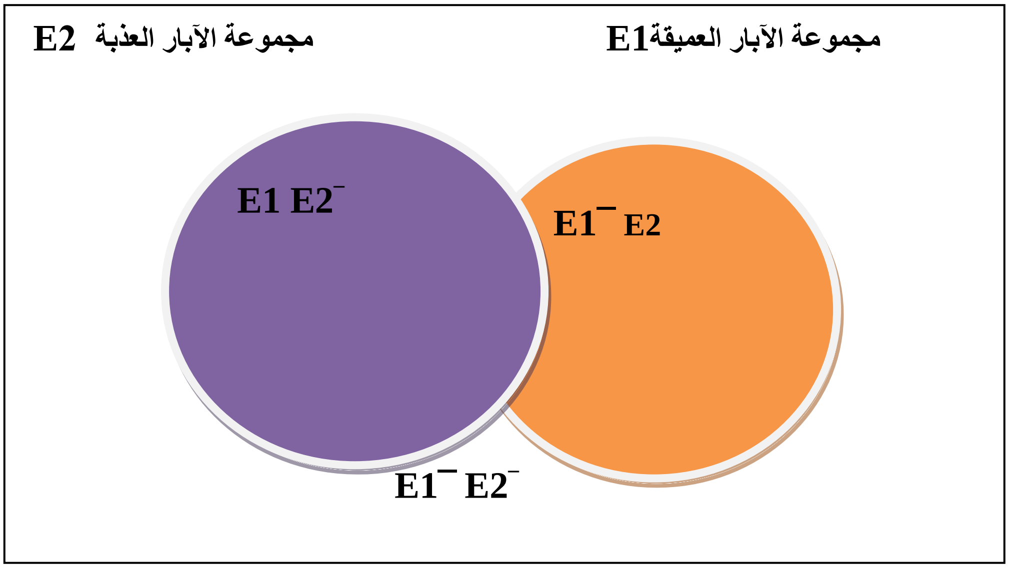
P (E) / $P(E^{-})$ =

(3/5) /(2/5)=3/ 2

مخطط فن : هو مخطط يستخدم لتوضيح موقع وعدد العناصر التي ترتبط مع كل حادثة نسبة الى المجتمع ،والذي يرمز له (u)ففي المثال السابق (u=150) ،لنفرض ان الحادثة E1 تمثل مجموعة الابار العميقة التي تتألف من 80% عذبة وان الحادثة E2 تمثل مجموعة الآبار العذبة ،ويمكن توضيح ذلك في مخطط فن كما مبين بالشكل رقم (1) : من مخطط فن بالشكل رقم (1) يمكن كتابة المعادلة:

$$\mathbf{n\,(E_i) + n\,(\bar{E}_i) = n(u)}\ .$$

$$\mathbf{P\,(E_i) = n\,(E_i)/n(u).}$$



$n\,(E\,i)$ =عدد الغناصر التي ترتبط بالحادثة($E\,i$) .
$n\,(u)$=عدد الغناصر التي ترتبط بالحادثة $E\,i$.اولا اي ترتبط بالمجتمع.
$n\,(\bar{E}\,i)$=العدد الكلي العناصر الموجودة بالمجتمع u غير مشمولة في $E\,i$.

للمثال اعلاه: $n\,(E_i) + n\,(\bar{E}\,i) = 75+75 = 150$

$$\mathbf{P\,(E1) = 75/150 = 0.5}$$

$$\mathbf{P\,(E2) = 90/150 = 0.6}$$

التقاطع : يرمز الى تقاطع الحادثة E1 مع الحادثة E2: (E1E2) اي ان جميع العناصر التي تقع E1 وE2 وفي آن واحد.لذا يكون.

$$\mathbf{P(E1E2) = [n\,(E1E2)]\,/\,[n\,(u)]}$$

بالنسبة للمثال اعلاه: $P(E1E2) = 60/150 = 0.4$

وهو احتمال حصول بئر عميق وعذب في آن واحد

الاتحاد : ان اتحاد الحادثة E1 مع الحادثة E2 : يشمل كل العناصر التي تقع في E1 و E2 او في كليهما.ويرمز له (E1&E2) اي :

$$\mathbf{E1\&E2 = E1+E2}$$

$$\mathbf{P(E1E2) = 105/150 = 0.7}$$

مثال : بامكان المهندس الاول انجاز تصميم بصوره مضبوطه باحتمال4/5،والمهندس الثاني باحتمال 3/4،والمهندس الثالث باحتمال 2/3.ماهو احتمال انجاز التصميم بصوره مضبوطه.

الحل:ان نسبة خطأ المهندسيين الثلاثة في آن واحد $P(E1.\bar{E}2.E3)=1/5*1/4*1/3=1/60$

وبذلك يكون احتمال انجاز التصميم بصورة مضبوطه $P(E1+E2+E3)=1-1/60=59/60$

الاحتمال الشرطي: يمثل الاحتمال الشرطي للحادثة E1 امكانية حدوثها شرط اصطحاب ظروف معينة فمثلا $P(E1/E2)$ يرمز لاحتمال E1 شرط ان تكون من مجموعة E2 وبذلك فان الاخيرة هي التي تمثل المجتمع،ويمكن توضيح الحالة في مخطط فن كما مبين بالشكل ويمكن التعبير عنه بدلالة المعادلة: $P(E1/E2)= P(E1/E2)/ P(E2)$

المجتمع E2

مخطط فن للاحتمال الشرطي

اي ان: $P(E1E2)= P(E1/E2)/ P(E2)= P(E2 / E1)= P(E1/E2) .P(E1)$

$$P(E1/E2)= [P(E2 / E1) .P. (E1)]/ P(E2)$$

$$P(E1/E2)+ P(\bar{E}1/E2)=1$$

مثال : في مخزن للاسمنت 1000كيس منها700 كيس مقاوم للاملاح و200 كيس منها غير صالح للاستعمال،وان نصف غير الصالح للاستعمال من النوع المقاوم للاملاح .عند سحب كيس واحد بصورة عشوائيه من المخزن جد:

- ✓ **احتمال ان يكون مقاوم للاملاح وغير صالح للاستخدام.**
- ✓ **احتمال ان يكون صالح للاستخدام اذا كان مقاوم للاملاح.**
- ✓ **احتمال ان يكون مقاوم للاملاح اذا كان غير صالح للاستخدام.**

الحل :

$P(E1)=200/1000=0.2$

$P(E2)=700/1000=0.7$

$P(E1E2)=100/1000=0.1$

1-$P(E1E2)=1/10$.

2-$P(E1/E2)= P(E1E2)/P(E2)=1/10*7/10= 1/7$

3-$P(E2/E1) = / P(E2E1)/P(E1)=1/10*2/10=0.5$.

الاحداث المستقلة **: يقال للحدثين E1&E2 مستقلتين حيث لاتؤثرنتيجة احدهما على نتيجة الآخرى .**

$P(E1/E2)= P(E1)$

$P(E1E2) = P(E1).P(E2)$

$P(E2/E1)=P(E2)$

$P(E1).P(E2) > 0$

الاحداث المتنافية: **لاتوجد في الاحداث المتنافية عناصر مشتركة وكما موضح بمخطط فن التالي:**

$P(E1+E2) = P(E1)+P(E2)$

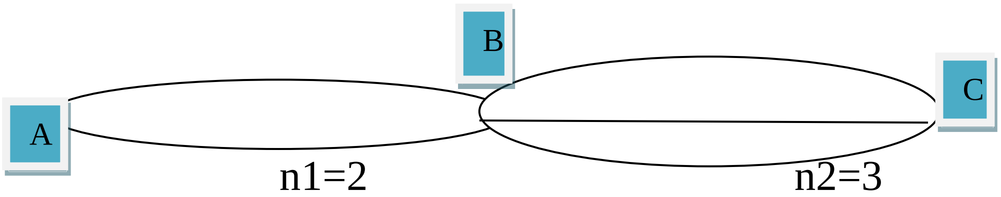
H.W3

في مجمع سكني 1000 شقة ،500في القاطع الشمالي وال 500 الاخرى في القاطع الجنوبي في كل قاطع 200 شقة من الشقق تحتوي على شبابيك كبيرة و100 مدفئة مركزيا،30% من الشقق ذات الشبابيك الكبيرة مدفئة مركزيا عند سحب شقة بصورة عشوائية جد احتمال كونها:

- ☒ في القاطع الشمالي
- ☒ في القاطع الشمالي وذات شبابيك كبيرة
- ☒ في القاطع الشمالي وذات شبابيك كبيرة ومدفئة مركزيا
- ☒ في القاطع الجنوبي وغيرمدفئة مركزيا

التبديلات والتوفيقات

قاعدة الضرب:عندما يراد اختيار واحد من العناصر n1 المصاحبة للحادثة E1 واختيار واحد من العناصر n2 المصاحبة للحادثة E2 فان هناك n1 n2 من الحالات المختلفة لحدوثهما معا. فمثلا للانتقال من A الى Cمرورا ب B يمكن اختيار احد الطريقتين بين A وB واحد الطرق الثلاثة بين B وC فيكون عدد المسارات الممكن اتخاذها بين A و C(2*3 =6) .



قاعدة الجمع :اذا كانت الحادثتان E1 وE2 *متنافيتين والاولى تحدث ب* n1 *من الطرق والثانيه ب* n2 من الطرق، اي ان عدد العناصر المصاحبة للاولى والثانية n1 و n2 على الترتيب فان واحد منهما او الاخرى تحدث ب n1+ n2 . من الطرق ،ويمكن تصميم قاعدة الجمع لتشمل اكثر من حادثين فمثلا اذا كان عدد انواع الحفارات السلكية 2 والهيدروليكية 3 فان هناك(3+2=5) من طرق الحفر المختلفة .

التبديلات:تعني التبديلة ترتيب جميع عناصر او جزء من عناصر مجموعه ،بحيث يكون للتسلسل اهمية فمثلا بالنسبه لترتيب مجموعه مؤلفة من ثلاث حروف A,B,C فان هناك ثلاث اختيارات للموقع الاول واثنين للموقع الثاني وواحد للموقع الثالث .فيكون عدد الترتيبات الممكنه :

3*2*1=6 ! 3 وهي:

ABC BAC CAB

ACB BCA CBA

ويمكن تعميمها الى اي جزء n فيكون عدد الترتيبات الممكنه ..

$$nPn = n!$$

nPn = تشير الى تباديل مجموعة مكونه من n من العناصر،اما اذا رتبت r من العناصر الماخوذه من من n من العناصر الكلية فيكو عدد الترتيبات الممكنة :

$$nPr = n(n-1)(n-2)\text{--------------}(n-r+1)$$

وبعد ضرب المعادلة ب $(n-r)!/(n-r)!$ تصبح: $nPr = n! / (n-r)!$

عندما r=n تاخذ المعادلة الصيغة التالية: **$nPr = n! / (n-r)! = n! / !0 = n!$**

مثال:اذا كان في صف التدريسي 50 طالبا وسمح لثلاثة فقط بتوجية الاسئلة ،جد الطرق المختلفة التي تأتي بها الاسئلة .

الحل: 50P3= 50!/(50-3)!=50*49*48=117.600

ملاحظة: عندما يتكرر العنصر في المجموعه n يقل عدد التبديلات ،فاذا تكرر عنصر واحد في هذه المجموعه اصبح عدد التبديلات نصف العدد بدون تكرار هذا العنصر.

- فمثلا اذا استبدلنا مجموع الحروف ABC بالمجموعة AAB تصبح تبديلات الاخيرة 3 وهي AAB,BAA, ABA

3! /2! =3

وبصورة عامة nPn=n!/(n1! n2!.......nk!)

حيث ان : n1,n2,.......nk تمثل عدد العناصر المتطابقة للمجموعة الاولى والثانية و.......الى المجموعه k والواقعة ضمن المجموعة الكلية n,اي: n1+n2+n3+...... nk= n

والمثال التالي يوضح هذه الحالة.

مثال:ماعدد تبديلات مجموعة الحروف AAABBCCDE

الحل :بما ان هناك 3A,2 B, 2C,1D,1E فيكون عدد التبديلات :

9!/(3!2!2!1!1!) = (9*8*7*6*5*4!)/24=15120.

التوفيقات :تمثل التوفيقات عدد المجاميع المختلفة الممكن تكوينها بحجم r من العناصر الكلية n دون النظر الى الترتيب ، ويرمز لها ب nCr لكل مجموعة من المجاميع هذا العدد(rPr=r!) من الترتيبات المختلفة،لذا يكون عدد المجاميع مساويا الى عدد التبديلات مقسوما على عدد التبديلات ضمن المجموعة الواحدة.

nCr= nPr/rPr =n! / [(n-r)! r!]

مثال:ماهو عدد المجموعات المخلتفة التي يمكن ان يتم بها اختيار ثلاثة طلاب من مجموع 50 طالبا ؟

الحل : يكون عدد المجاميع المختلفة مساويا الى

50C3=50! / (3! 47!)=117600/6=19600

H.W4

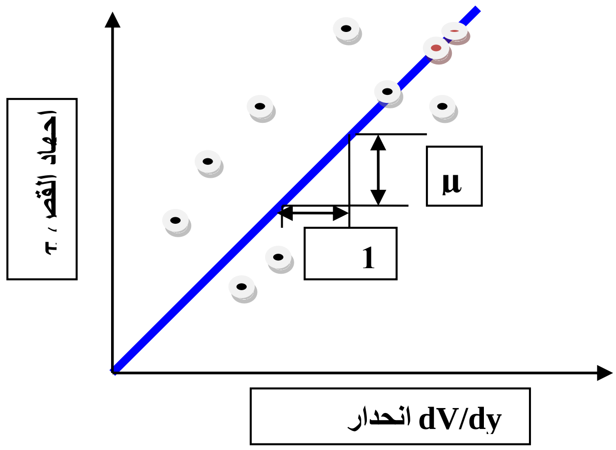
في النية اختيار لجنة طلابية مؤلفة من 5 طلاب فاذا كان في الصف 40طالبا و20 طالبة ماهو احتمال ان تكون اللجنه مؤلفة من طالبتين وثلاثة طلاب اذا تم الاختيار بصورة عشوائية ؟؟

الانحدار والارتباط: نحتاج في كثير من الامور معرفة قيمة المتغير التابع (Y) بدلالة المتغير المستقل (X) فيعبر عن العلاقة بين هذين المتغيرين بالمعادلة :

$$Y = f(x)$$

اي ان Y تابع ل X فمثلا المعادلة: $\tau = \mu(dV/dy)$

تعبر عن العلاقة الخطية بين اجهاد القص(τ) وميل السرعه (dV/dy) في الجريان الطباقي حيث يكون ميل الخط مساويا الى الثابت μ الذي يمثل اللزوجة .هنا يكون اجهاد القص المتغير التابع وانحدار السرعة المتغير المستقل .لكن عند القيام بتجربة مختبرية نجد عدم ترتيب النقاط (μ,dV/dy) على خط مستقيم بل تكون مبعثرة وتشكل مايسمى بمخطط الانتشار .



مخطط الانتشار للقيم (μ,dV/dy)

اختيار المنحنيات: عندما تتوفر البيانات عن المتغير التابع Yوالمتغير المستقل X يصبح بالامكان رسمها على الورقة البيانية فنحصل على مايسمى بمخطط الانتشار الذي يوضح العلاقة او الترابط بصورة اولية .فيمكن رسم منحنى تقريبي للنقاط بمجرد النظر وتكون الطريقة ملائمة اذا كانت النقاط قريبة للمنحني ،لذلك فان الخطوة الاولى هي الحصول على منحنيات تقريبية ملائمة للبيانات تحت الدراسة.

هناك عدة منحنيات تقريبية معروفة ابسطها منحنى الدرجة الاولى (الخط المستقيم) الذي يعبر عنه بالمعادلة

$$\hat{Y} = a + bX$$

حيث ان $\hat{Y}$ تمثل قيم التي يتم الحصول عليها من هذه المعادلة بعد تعويض قيم Xالمشاهدة .

المنحني من الدرجة الثانية يسمى بالقطع المكافئ حيث تاخذ معادلته الصيغة :

$$\hat{Y} = a + bx + cX^2$$

وتكون معادلة المنحنى من الدرجة n بالصيغة :

$$\hat{Y} = a_o + a_1x + \dots\dots\dots + a_{n-1}X^{n-1} + a^nX^n$$

اذ يسمى الجزء الايمن من هذه المعادلة بعد جعله مساويا للصفر بمتعدد الحدود .ويمكن الاستفادة من مختلف الاوراق البيانية مثل الورقة الحسابية والورقة اللوغارتمية و الورقة اللوغارتمية – الحسابية في ايجاد

صيغة المعادلة بصورة سريعة وملائمة، فاذا اقتربت البيانات من الصيغة الخطية على الورقة اللوغارتمية تكون صيغة المعادلة الملائمة:

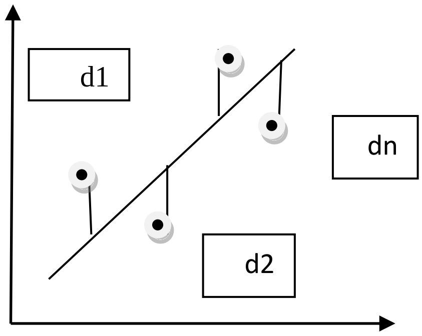
$$\text{Log}\hat{Y}=\text{Log}a+b\ \text{Log}X \qquad \text{او} \qquad \hat{Y}=a\ x^{b}$$

حيث ان a قيمة التقاطع و b ميل الخط ،اما اذا اقتربت البيانات من الصيغة الخطية على الورقة اللوغارتمية – الحسابية فتكون الصيغة الملائمة للمعادلة :

$$\text{Log}\hat{Y}=\text{Log}a+bX \qquad \text{او} \qquad \hat{Y}=a10^{b\,x}$$

طريقة المربعات الصغرى : عندما تتوفر البيانات بالنسبة الى قيم المتغير المستقل X والقيم المرادفة لها بالنسبة للمتغير التابع Y يصبح من الممكن رسم العلاقات .حيث يتم اختيار المنحنى بطريقة لمربعات الصغرى على اساس ان مجموع مربعات المسافات الرأسية للنقاط في مخطط الانتشار عن المنحنى في حده الادنى كما موضح بالشكل ادناه فبالنسبة للعلاقة الخطية .

$$\sum_{i=1}^{n} \mathbf{di}\hat{}2=\ \mathbf{d^2}=\sum_{i=1}^{n}=\ (\hat{\mathbf{Y}}\mathbf{i}-\mathbf{Yi})2$$



حيث تمثل Y القيم المشاهدة و$\hat{Y}$ القيم المحسوبة من المعادلة الناتجه من جودة الموافقة وان n تمثل عدد ازواج البيانات ،يمكن التعبير عن:

$$d^{2=}\ |(a+bx1)-Y1|+^{=}|(a+bx2)-Y2|+........+^{=}|(a+bxn)-Yn|$$

بتفاضل هذه المعادلة مع a وb نحصل على مايسمى بالمعادلات الطبيعية :

$$\sum Y_{\,i\,=}\ na\ _{+}b\sum X_i$$

$$\sum\ X_i\ Y_i = a\sum Xi + b\sum X_i^{2}$$

وبحلها نحصل على

$$a=\sum_{Yi}\sum X_i^{\ 2}\ -\sum X_i\sum\ X_i\ Y_i\ \ /n\sum X_i^{\ 2}-(\sum X_i\,)^2$$

$$b = n\ \sum\ X_i\ Y_i\ \ -\sum X_i\sum Y_i\ \ /n\sum X_i^{\ 2}-(\sum X_i\,)^2$$

ويمكن ايجاد قيمة a من المعادلة التالية :

$$a= \bar{Y}-b\,\bar{X}=\sum Yi-b\sum\ Xi/\ n$$

لملائمتها الحسابات حيث ان $\bar{X}$ و $\bar{Y}$ تمثلان وسط قيم Xi و Yiعلى الترتيب.وبعد الحصول على قيمتي a و b نحصل على اجود معادلة خطية التي تعرف ايضا بمعادلة الانحدار ل Y وX ويسمى

$$\boxed{\hat{Y}=a+bX}$$

الخط الناتج من رسم هذه المعادلة بخط الانحدار.وعلية فان معادلة الانحدار لمعادلة من الدرجه الاولى تكون بالصيغه التالية

اما معادلة الانحدار لمعادلة من الدرجه الثانية فتكون بالصيغه التالية:

$$\hat{Y}=a+bx+cx^2$$

- **ملاحظة:** يمكن الحصول على ثوابت اجود معادلة خطية بالاسلوب نفسه اعلاه ففي حالة اعتبار Y المتغيير المستقل للحصول على اجود معادلة خطية تكون الصيغة:

$$X\hat{}= \bar{a}+\bar{b}Y$$

ويتم الحصول على قيم $\bar{a}$ و $\bar{b}$ بنفس الطريقة التي استخدمت لحساب a,b فيكون مجموع مربعات المسافات الافقية (بدلا من الرأسية) عن خط الانحدار في مخطط الانتشار :

$$d\hat{}2 = \sum_{i=1}^{n} = (X\hat{} - Xi)\hat{}2$$

- **ملاحظة** : يمكن معرفة جودة الموافقة لاي معادلة باستخدام اختبار كاي تربيع بمستوى دلالة معين باعتبار $\hat{Y}$ القيم النظرية (المتوقعه) و **Y** القيم الميدانية(المشاهدة).

مثال: جد اجود معادلة خطية للبيانات ادناه ،اختبر جودة الموافقة باستخدام اختبار مربع كاي بمستوى دلالة $\alpha = 0.05$ ؟

X	1	2	3	4	5
Y	2	3	5	8	9

الحل: نجري الحسابات كما مبين بالجدول التالي :

X	Y	X^2	Y^2	XY	$\hat{Y}$	$(\hat{Y}-Y)^2/\hat{Y}$
1	2	1	4	2	1.6	0.1
2	3	4	9	6	3.5	0.07
3	5	9	25	15	5.4	0.03
4	8	16	64	32	7.3	0.07
5	9	25	81	4	9.2	0.00
15	27	55	183	100		0.27

$\bar{X}=\sum X/n=15/5 = 3.$

$\bar{Y} = \sum Y/n = 27/5 = 5.4.$

$$\mathbf{a = \sum_{Yi} \sum X_i^{\,2} - \sum X_i \sum X_i Y_i \;/n \sum X_i^{\,2} - (\sum X_i)^2}$$

$$a = 27*55 - 15*100/5*(55) - (15)^{2} = -0.30.$$

$$\mathbf{b = n \sum X_i Y_i - \sum X_i \sum Y_i \;/n \sum X_i^{\,2} - (\sum X_i)^2}$$

$$b = (5*100)-(15*27)/(5*55)-(15)^2 = 1.90.$$

فتكون معادلة الانحدار: $\underline{\hat{Y} = -\ 0.03 + 1.90\,X}$

لاختبار درجة الموافقة نختار بين الفرضيتين :

1-الفرضية الصفرية لافرق بين القيم المشاهدة والقيم النظرية المتوقعه .

2-الفرضية المتناوبة H_1 وجود فرق بين القيم المشاهدة والقيم النظرية المتوقعه.

ومن اجل ذلك نحسب الاحصائية :

$$\chi = \sum (Y-\hat{Y})^2/\hat{Y} = 0.27$$

$$^{2}\chi_{=}[0.95, 4] = 9.49$$

وهي اقل من القيمة الجدولية الحرجه لذا نقبل H_0 .

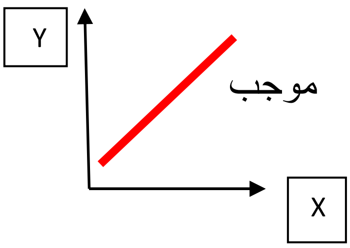
نجد $\bar{a}$ و $\bar{b}$:

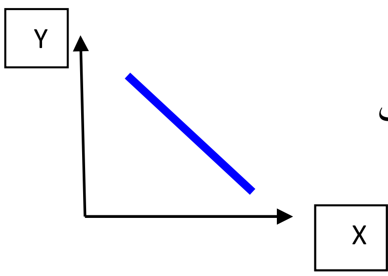
$\mathbf{\bar{a} = 0.24}$

$\mathbf{\bar{b} = 0.51}$

$\mathbf{\hat{X} = 0.24 + 0.51Y.}$

الترابط: نحتاج في كثير من المسائل الهندسية والدراسات المختبرية والميدانية الى معرفة مقدار الترابط بين المتغير التابع والمستقل ففي التجارب المختبرية برسم مخطط الانتشار على المستوى XY بين انحدار السرعة والاجهاد او بين عمق الماء وسرعته هكذا . فقد يكون الترابط كاملا بين المتغيرين كما يحصل بين نصف القطر وحجم الكرة وقد لايكون كاملا كما يحصل بين عمق الماء والسرعة واحيانا لايوجد اي ترابط اذا كان المتغير المستقل لايؤثر على المتغير التابع. يسمى الترابط بين متغيرين اثنين ب (**الترابط البسيط**) فاذا كانت المعادلة المعبرة خطية سمي بالترابط الخطي البسيط او الترابط الخطي. فقد يكون الترابط سالبا او موجبا. كما مبين بالشكل ادناه:





سالب

معامل الترابط الخطي : يعبر عن مدى الترابط بين المتغيرين ويعرف بالمعادلة التالية:

$$r = [\sum (X_i - \bar{X})(Y_i - \bar{Y})] / [(n-1)\, SX\, SY]$$

حيث ان:

$\bar{X}$& $\bar{Y}$= وسط قيم X_i، Y_i على الترتيب.

SX،SY = الانحراف المعياري لقيم X_i، Y_i على الترتيب.

- **اذاكان الترابط بين X و Y كاملا وموجبا اصبحت (r =1) اي ان العلاقة بين X و Y علاقة طردية .**
- **اذاكان الترابط بين X و Y كاملا وسالبا اصبحت (r = - 1) اي ان العلاقة بين X و Y علاقة عكسية.**
- **في حالة عدم اعتماد X على Y اصبحت(r = 0) .**
- **في حالة $-1 < r < 1$ فيكون الترابط غير كامل ويحتاج الى اختبار الدلالة.**
- **يطلق على معامل الترابط r بمعامل بيرسون العزومي للترابط.**

للحسابات اليدوية نستخدم المعادلة التالية :

$$r = \frac{n(\sum XiYi) - (\sum Xi)(\sum Yi)}{\sqrt{[n(\sum Xi^2) - (\sum Xi)^2]}\ \sqrt{[n(\sum Yi^2) - (\sum Yi)^2]}}$$

ويمكن ايجاد الخطأ المعياري للتخمين او انحراف العينه المعياري للانحدار من العلاقة التالية :

$$SX.Y = \left[\sqrt{\sum (Yi - \hat{Y})^2 / (n-2)}\right]$$

مثال :للبيانات في الجدول ادناه جد معادلة الانحدار واحسب معامل الترابط الخطي و الخطأ المعياري للتخمين ؟

X	1	2	3	4	5	6	7
Y	1	3	6	8	12	15	20

يمكن اجراء الحسابات كما في الجدول ادناه:

X	Y	X^2	X Y	Y^2	$\hat{Y}$	$Y-\hat{Y}$	$(Y-\hat{Y})^2$
1	1	1	1	1	-0.03	1.03	1.06
2	3	4	6	9	3.08	-0.08	0.01
3	6	9	18	36	6.19	-0.19	0.04
4	8	16	32	64	9.30	-1.3	1.69
5	12	25	60	144	12.41	-0.41	0.17
6	15	36	90	225	15.52	-0.52	0.26
7	20	49	140	400	18.63	1.37	1.88
∑ 28	65	140	347				5.11

a=∑Yi ∑Xi 2 - ∑ Xi ∑ Xi Yi /n ∑ Xi 2 -(∑ Xi)2= -3.14

$a = \sum Y_i \sum X_i^2 - \sum X_i \sum X_i Y_i / n \sum X_i^2 - (\sum X_i)^2 = -3.14$

$b = n \sum X_i Y_i - \sum X_i \sum Y_i / n \sum X_i^2 - (\sum X_i)^2 = 3.11$

معادلة الانحدار :

$\hat{Y} = -3.14 + 3.11X$

معامل الترابط الخطي:

$$r = \frac{n(\sum XiYi) - (\sum Xi)(\sum Yi)}{\sqrt{[n(\sum Xi^2) - (\sum Xi)^2]}\ \sqrt{[n(\sum Yi^2) - (\sum Yi)^2]}}$$

$r = (7)(347)-(28)(65)/\sqrt{[(7)(140)-(28)^2]}\sqrt{[(7)(879)-(65)^2]} = \underline{0.99}$

و الخطأ المعياري للتخمين :

$SX.Y = [\sqrt{\sum (Yi-\hat{Y})^2 / (n-2)}] = \sqrt{(5.11/5)} = \underline{1.01}$

معامل الترابط الترتيبي : في بعض المسائل الهند سية يكون من الصعب تحديد قيم للمتغيرين التابع والمستقل لذا توضع صفة ترتيب وتدرج بالنسبة للمتغيرين .لمعرفة مدى تأثير احد المتغيرين على الآخر نستخدم معامل الترابط الترتيبي ولقد اقترح سبيرمان المعادلة التالية:

$$r_s = 1 - 6 \sum_{i=1.0}^{n} di^2 / n(n^2 - 1)$$

ويسمى ب معامل الترابط الترتيبي لسبيرمان . وتستخدم هذه المعادلة في الحالات التي يكون فيها n بين 25 &30 او قل

di = الفرق بين ترتيب Xi &Yi .

<u>مثال</u> : في مسح شامل للتربة وبيان مدى تأثير كمية المطر على انتفاخها ،فقد استخدم اسلوب الترتيب فكانت النتائج كما موضح في الجدول ،احسب معامل الترابط الترتيبي لهذه البيانات؟

ترتيب كمية المطر X	1 2 3 4 5 6 7 8 9 10
ترتيب الانتفاخ Y	2 1 4 3 5 6 8 7 10 6

الحل: يمكن اجراء الحسابات كما في الجدول ادناه

X	Y	d	d^2
1	2	-1	1
2	1	1	1
3	4	-1	1
4	3	1	1
5	5	0	0
6	6	0	0
7	8	-1	1
8	7	1	1
9	10	-1	1
10	9	1	1

$$r_s = 1\text{-}6 \sum_{i=1.0}^{n} di^2 / n(n^2 - 1)$$

$$r_s = 1\text{-}6*8/10(100\text{-}1) = 0.952$$

HW5

في دراسة لمعرفة مدى تأثير ارتفاع البنايات على الشقوق استخدم اسلوب الترتيب فكانت النتائج كما موضح في الجدول ،احسب معادلة الانحدار ، معامل الترابط الخطي ، الخطأ المعياري للتخمين ومعامل الترابط الترتيبي لهذه البيانات؟

X ترتيب الارتفاع	1 2 3 4 5 6 7 8 9 10
Y ترتيب الشقوق	2 1 4 3 6 4 6 9 7 8