الاحصاء الهندسي

تعاريف ومصطلحات:

1-الاحصاء: وهو أحد أهم فروع الرياضيات وهو من أهم الفروع وهو يستخدم في كل مجالات الحياة. كما ويدخل علم الإحصاء في كل العلوم الأخرى وكل العلوم مهما كان تنوعها تحتاج إلى علم الإحصاء فمثلا :(العلوم العلمية – العلوم الإنسانية والأدبية – العلوم الرياضية – الفنون –العلوم العسكرية – العلوم الاجتماعية) لذلك يعتبر علم الإحصاء هو الأساس في كل العلوم.

•ما هي وظيفة علم الإحصاء(بماذا يهتم علم الإحصاء):

يهتم علم الإحصاء بشكل أساسي في جمع وتنظيم وتلخيص وعرض وتحليل وتفسير ونشر البيانات الإحصائية.

2-البيانات:

هي جزء من المعلومات الموثوق والمدعم بالأرقام أو معلومات تكون مرتبة ومنظمة ومخزنة بطريقة سهلة التداول و الفهم والتفسير ومدعمة بالأرقام ويمكن تحليلها وتفسيرها وليست كل المعلومات بيانات.

لماذا يعتبر علم الإحصاء ضروري لكل العلوم؟:

كل علم من العلوم فيه بيانات ويتعامل مع البيانات وبالتالي نحتاج إلى الإحصاء ليقوم بمعالجة وتحليل وتقييم وتفسير البيانات . أو بتعبير أخر كل العلوم أساسها البيانات وهي تحتاج إلى د ا رسات وأبحاث تطبيقية واختبارات وكل بحث أو اختبار أو تجربة يحتاج إلى علم الإحصاء.

من أين نحصل على البيانات:؟

١ - تأتي البيانات بشكل أساسي من مراقبة الظواهر ومشاهدتها في الطبيعة.

٢ - من خلال إجراء التجارب والاختبارات.

٣ -من خلال الأبحاث والد راسات والمقابلات.

فوائد علم الاحصاء:

١ -يمكن تحليل ومعالجة البيانات.
٢ -تفسير النتائج للأبحاث والتجارب والاختبارات.
٣ - الكشف عن صلاحية وجودة البيانات لأي بحث ولأي منتج ولأي تجربة جديدة.
٤ -تقييم أو تقويم الأبحاث والدراسات.

3-المجتمع:

عدد لانهائي من الإفراد أو العناصر التي تتعايش مع بعضها البعض وتتميز بخصائص ومواصفات تميزها عن بقية المجتمعات حيث أنه لكل مجتمع خصائص ومميزات تميزه عن المجتمعات الأخرى.

المجتمع البشري يقسم إلى مجتمعات صغيرة (إحصائية) هذا المجتمع الإحصائي معروف عدده وتميزه مجموعة من الصفات الزمانية و المكانية وهذا النوع هو الذي يستخدم في دراسات الأبحاث.

اهم ميزة في المجتمع الاحصائي عدد الافراد (n) و المتوسط (X).

4-العينة:

هي جزء من المجتمع يجب ألا يقل عدد أفرادها عن (٢- ١٠%) من عدد افراد المجتمع وهي تؤخذ بطريقة عشوائية وتوجد منها الأنواع التالية

أ- العينة العشوائية البسيطة: هي أصغر وأبسط أنواع العينات تؤخذ بطريقة عشوائية ويجب أن لا يقل عدد أفرادها عن(٣٠).

ب- العينة العشوائية الطبقية: وهي تؤخذ من المجتمع إذا كان على شكل طبقات أو إذا كان الهدف هو دراسة طبقات المجتمع. مثلا (طبقة الأثرياء – طبقة الأغنياء – طبقة الوسط – طبقة الفقراء – طبقة المعدمين).

ت- العينة العشوائية المنتظمة: تؤخذ بطريقة عشوائية ولكن في فترات منتظمة أو من أماكن منتظمة. مثلاً شركة تنتج مادة تنظيف وأردنا اختبار المنتجات فنقوم مثلا بأخذ العينات بفترات زمنية محددة (كل ساعة او يوم او أسبوع.....).

ث- العينة العشوائية العنقودية: تؤخذ إذا كان المجتمع مرتب بطريقة عنقودية أو كان على شكل تنظيمات هرمية . مثل التنظيم الإداري للاماكن في العراق (محافظة- قضاء-ناحية-قرية ....وهكذا) او (جامعة-كلية-قسم –مرحلة-شعبة-مجموعة)

٥- **المتغيرات :** هي توابع ذات قيم متغيرة دائما بصفة معينة وهي على نوعين

أ- **المتغيرات الكمية:** هي متغيرات ممكن قياسها بالمقاييس المعروفة كالطول و الوزن والحجم ودرجة الحرارة ويمكن التحكم بها وهي عادة ما تكون متسلسلة ومستمرة كما انها عادة ما تكون اعداد صحيحة او غير صحية (كسرية). كما انها يمكن ان تكون غير مستمرة كعدد البيوت الصغيرة في مكان ما او عدد الأسواق في مدينة ما او كمية مياه الامطار المتجمعة في وقت ما والتي دائما تمثل بأرقام صحيحة....وهكذا.

ب- **المتغيرات النوعية:** هي متغيرات لا يمكن قياسها بالطرق المعروفة مثل تدرجات اللون او درجة ذكاء.

كما يمكن تصنيف المتغيرات الى :

أ- **متغيرات تابعة:** وهي تلك المتغيرات التي نحاول تفسيرها ومعرفة أسباب حدوثها وتحديد مدى إمكانية التنبؤ بها والتي عادة ما تعتمد في تغيرها على المتغيرات المستقلة.

ب- **متغيرات مستقلة:** وهي التي تلعب دوراً مباشر في حدوث المتغيرات التابعة وعادة ما تستخدم لتفسير وفهم أسباب التغييرات على تلك تالك المتغيرات.

إضافة الى التقسيمين اعلاه يمكن ان تقسم المتغيرات الى عشوائية (درجة الحرارة ونسبة الرطوبة و سرعة الرياح ..وغيرها) و لا عشوائية او منتظمة (كمية العلاج المأخوذ او السماد المضاف او عدد مرات الري او نسبة حبيبات النانو المضافة الى سائل معين.

٦- الصفة المدروسة (Category):
هي الصفة التي يختارها الباحث ليدرس تغيراتها اثناء فترة الدراسة او البحث. مثل (الكثافة ، الموصلية، الكفاءة، الطاقة الخارجة، معدل صرف الوقود ..وما الى ذلك).

٧- القيمة الإحصائية : هي القياس او الرقم الذي نحصل عليه عند قياس الصفة المدروسة ومجموع القيم الإحصائية نسميها البيانات.

٨- التوزيع: هو طريقة توزيع القيم الإحصائية او توزيع البيانات للمجتمع او لعينة واحتمالاتها

الرموز الإحصائية :
اذا كان لدينا مجموعة بيانات كالاتي y={2,5,9,13,20,38}
فان y1=2, y2=5,……y6=38
واذا كان هناك مجموعة أخرى من البيانات تختلف عن الأولى فيمكن ان نرمز لها برمز مختلف كان يكون x

x={9,20,27,1,0,7}
أي ان x1=9, x2=20,……x6=7

كما ان مجموع عناصر المجموعة الأولى (y) يمكن ان تكون بالصيغة الاتية $\sum_{i=1}^{n} yi = y1 + y2 + \cdots + yn - 1 + yn$

كذلك يمكن جمع مربع عناصر المجموعة

$$\sum_{i=1}^{n} y_i^2 = y_1^2 + y_2^2 + \cdots + y_{n-1}^2 + y_n^2$$

كذلك

$$(\sum_{i=1}^{n} y_i)^2 = (y_1 + y_2 + \cdots + y_{n-1} + y_n)^2$$

الاحصاء الهندسي

$$\sum_{i=1}^{n} x_i * y_i = x_1 * y_1 + \cdots + x_n * y_n$$

$$(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i) = (x_1 + \cdots + x_n)(y_1 + \cdots + y_n)$$

$$\sum_{i=1}^{n} C\, y_i = C \sum_{i=1}^{n} y_i$$

$$\sum_{i=1}^{n} x_i + y_i = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$$

مثال (1): اذا توفرت لدينا بيانات المتغيرين الآتيين؛

y= 3, 9, 6, 2 ; x= 4, 2, 3, 7

استخدم قوانين جمع قيمهما

الحل:

$\sum_{i=1}^{n} y_i = 3 + 9 + 6 + 2 = 20$

$\sum_{i=1}^{n} y_i^2 = 3^2 + 9^2 + 6^2 + 2^2 = 130$

$(\sum_{i=1}^{n} y_i)^2 = (3 + 9 + 6 + 2)^2 = 20^2 = 400$

$\sum_{i=1}^{n} x_i * y_i = 4 * 3 + 2 * 9 + 3 * 6 + 7 * 2 = 62$

$(\sum_{i=1}^{n} x_i)(\sum_{i=1}^{n} y_i) = (4 + 2 + 3 + 7)(3 + 9 + 6 + 2) = 16 * 20 = 320$

مثال (2) : اذا توفرت لدينا بيانات المتغيرين الآتيين؛

y= 3, 9, 6, 2 ; x= 2, 6, 3, 1

استخدم قواعد الجمع معها.

الحل:

i. $\sum_{i=1}^{n}(y_i - x_i)^2 = (3-2)^2 + (9-6)^2 + (6-3)^2 + (2-1)^2 = 20$
ii. $\sum_{i=1}^{n}(x_i - 3) * (y_i - 5) = (2-3) * (3-5) + \cdots + (1-3) * (2-5) = 20$
iii. $\sum_{i=1}^{n} x_i * y_i^2 = 2 * 3^2 + \cdots + 1 * 2^2 = 616$

الاحصاء الهندسي

جداول التوزيعات التكرارية: Frequency Distributions

تأخذ الظواهر الإحصائية المدروسة قيماً عددية كثيرة ومتكررة وفي بعض الأحيان تكون النتائج الملاحظة غير عددية ، فيمكن في هذه الحالات تحويلها إلى قيم عددية ، مثلاً يمكن تحويل " نعم " أو " لا " أو " صح " أو " خطأ " إلى " مع " أو " ضد " وبالتالي إلى " 1 " أو " صفر " ، مما يسمح لنا بتشكيل جداول تكرارية.

إن تصنيف وتبويب مجمل البيانات المدروسة يعني بالضرورة ترتيب هذه البيانات تصاعديا أو تنازلياً مما يسمح لنا استخلاص صورة واضحة عن المدى "Range" الذي تتراوح فيه البيانات على عدد من الفئات "Classes" معتبرين هذه الفئات وجوها للظاهرة المدروسة حيث يتم تفريغ المعلومات على أساس هذه الفئات، ومن ثم نحدد العدد المقابل لكل فئة من هذه الفئات لنستنتج تكرارات القيم العددية ضمن فئاتها . ونسمي الجدول الذي يضم الفئات والتكرارات المقابلة لها جدول التوزيع التكراري Frequency Distribution Table .

مثال (3):

تمثل القراءات التالية الفترات الزمنية (بالساعات) التي يقضيها 80 طالباً من طلاب كليات الجامعة في المكتبة أسبوعياً.

شكل جدول توزيع تكراري لهذه البيانات .

24	15	22	28	17	12	20	16	23	16
22	15	23	18	11	21	17	30	16	29
19	39	19	18	14	20	28	18	29	24
20	22	34	12	29	25	17	23	20	18
24	14	32	27	18	21	15	19	17	16
21	19	23	26	24	13	23	15	25	22
25	16	18	23	20	10	22	15	18	16
15	26	17	20	21	19	20	21	27	31

الحل :

بما أن أصغر قراءة هي 10 وأكبرها 39 فيمكن حساب **المدى** وهو الفرق بين القراءتين الكبرى والصغرى وهو R = 39-10=29 .

باختيار عدد الفئات 6 يمكن أن نحسب طول الفئة وهي هنا (29 ÷ 6 = 4.833 ≈ 5)

وفي مثل هذه الحالات يجب ان يقرب الناتج (عدد الفئات) للأعلى دوماً ، أي لأقرب عدد صحيح تالي. وبالتالي فإن حدود الفئات هي

10-14 , 15-19 , 20-24 , 25-29 , 30-34 and 35-39

لاحظ أنه يمكن اختيار عدد آخر للفئات فإذا اخترنا العدد 8 مثلاً فإن أطوال الفئات سيكون
(4 ≈ 3.625 = 8÷ 29) وسوف تكون الفئات الثمان كالتالي:

10-13 , 14-17 , 18-21 , 22-25 , 26-29 , 30-33 , 34-37 , 38-41

وهنا نلاحظ بأن الرقم 41 غير موجود في البيانات أصلاً ووجوده هنا لا يؤثر على البيانات الأصلية وإنما وضع لكي تكون أطوال الفئات متساوية كما ويمكن أن نبدأ بعدد أصغر من الرقم 10 الذي هو أصغر قراءة موجودة ولنفس السبب . لقد ذكرنا هنا إمكانيتين للتوضيح فقط وسنتابع شرح المثال على ست فئات والسبب في ذلك التفاصيل التي ذكرت سابقاً وهي أن أطوال الفئات يمكن أن تكون .... 5 , 10 , 100 ، ثم نصنف هذه البيانات وفق الفئات الست كما يأتي:

الفئات "ساعة"	إشارة tally	تكرار Frequency
10-14	𝍸 III	8
15-19	𝍸 𝍸 𝍸 𝍸 𝍸 III	28
20-24	𝍸 𝍸 𝍸 𝍸 𝍸 II	27
25-29	𝍸 𝍸 II	12
30-34	IIII	4
35-39	I	1

الاحصاء الهندسي

لاحظ في الجدول أن الحدود الفعلية للفئة الأولى 9.5-14.5 وللفئة الثانية 14.5-19.5 ..... وهكذا. ونحصل على هذه الحدود بطرح و جمع نصف وحدة الدقة للحد الادنى والحد الاعلى على الترتيب .

وفي الجدول أعلاه يمثل العمود في أقصى اليمين التكرار(Frequency) عدد مرات وقوع القراءات في الفئات الملائمة وهو ما ندعوه تردد الفئة أو **التكرار Class Frequency** .

إن أصغر رقم وأكبر رقم في عمود أقصى اليسار " الساعات hours " يعبر عنه **بحدود الفئات** Class limits فالأرقام

10 15 20 25 30 35

هي الحدود الدنيا للفئات بينما

39 34 29 24 19 14

فتسمى **الحدود العليا للفئات** Lower class limits and Upper class limits على الترتيب .

في هذا المثال قربت فترات الدراسة لأقرب ساعة وهذا يعني أن الفئة الأولى تغطي الفترة من 9.5 إلى 14.5 والفترة الثانية من 14.5 إلى 19.5 ساعة وهكذا ... لتدعى هذه الأرقام **حدود الفئات الحقيقية** "Class boundaries"أو " Real class limits"، إن حدود الفئات هي الأكثر استعمالاً أو شيوعاً.

تعتبر **مراكز الفئات Class marks** إحدى عناصر التوزيعات التكرارية الهامة وهي بالتعريف منتصف طول الفئة ونحصل عليه بإضافة الحد الأدنى للفئة لحدها الأعلى وقسمة الناتج على 2 أي (a+b 2 /(حيث a الحد الأدنى b الحد الأعلى . أما **طول الفئة Class interval** فهو الفرق بين الحد الأعلى الفعلي و الحد الأدنى الفعلي وغالباً ما يسمى مدى الفئة . نحصل على مراكز الفئات بحساب مركز الفئة الأولى ومن ثم نضيف له طول الفئة فنحصل على مركز الفئة الثانية ، نضيف طول الفئة للناتج فنحصل على مركز الفئة التالي... وهكذا . وبالعكس فإن الفرق بين مركزي فئتين متتاليتين يعطي طول الفئة .

قد نكون بحاجة إلى معيار آخر ويدعى **بتكرار التوزيع المئوي** Percentage-distribution ويحسب بقسمة تكرار كل فئة على المجموع العام وضرب الناتج بـ 100. أو معيارا آخر و يدعى **التكرار النسبي** والذي يحسب بقسمة التكرار لكل فئة على المجموع العام (Relative frequency). وهكذا يمكن تلخيص ذلك بجدول يحتوي نتائج المثال الحالي والسابق بالشكل الآتي:

مركز الفئة الحقيقية		التكرار	الفائات
١٢		٨	١٤-١٠
١٧		٢٨	١٩-١٥
٢٢		٢٧	٢٤-٢٠
١٢		١٢	٢٩-٢٥
١٢		٤	٣٤-٣٠
١٢		١	٣٩-٣٥

التكرار المئوي	التكرار النسبي	التكرار	مركز الفئة	حدود الفئات الحقيقية	الفائات
١٠	٠,١	٨	١٢	١٤,٥-٩,٥	١٤-١٠
٣٥	٠,٣٥	٢٨	١٧	١٩,٥-١٤,٥	١٩-١٥
٣٣,٧٥	٠,٣٣٧٥	٢٧	٢٢	٢٤,٥-١٩,٥	٢٤-٢٠
١٥	٠,١٥	١٢	٢٧	٢٩,٥-٢٤,٥	٢٩-٢٥
٥	٠,٠٥	٤	٣٢	٣٤,٥-٢٩,٥	٣٤-٣٠
١,٢٥	٠,٠١٢٥	١	٣٧	٣٩,٥-٣٤,٥	٣٩-٣٥
١٠٠	١	٨٠			الجموع

Cumulative Frequency
877

الفئات في حالة التراكمي الصاعد او المتزايد	الفئات	مركز الفئة	التكرار	التكرار التجميعي الصاعد	الفئات في حالة التراكمي النازل او المتناقص	التكرار التجميعي النازل
١٠ <			٠	٠	٤٠ <	80
١٥ <	١٤-١٠	١٢	٨	٨	٣٥ <	79
٢٠ <	١٩-١٥	١٧	٢٨	٣٦	٣٥ <	75
٢٥ <	٢٤-٢٠	٢٢	٢٧	٦٣	٣٠ <	63
٣٠ <	٢٩-٢٥	٢٧	١٢	٧٥	٢٠ <	36
٣٥ <	٣٤-٣٠	٣٢	٤	٧٩	١٥ <	8
٤٠<	٣٩-٣٥	٣٧	١	٨٠	١٠ <	0