مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ ـ ١٤٣١هـ

المحاضرة الاولي

الباب الاول

عناصر المحاضرة

1-1 مقدمة.

2-1 تعريف علم الاحصاء.

3-1 الهدف من علم الاحصاء.

١-١ مقدمة

اشتقت كلمة الإحصاء من اللفظ اللاتيني " **ستاتوس** – Status" بمعنى **الدولة** . وقد استعمل علم الإحصاء قديما استعمالات مبكرة تضمنت تجميع البيانات والتخطيط ووصف مظاهرة متعددة للدولة.

في عام 1662 قام العالم " **جون جرونت** " بنشر معلومات إحصائية حول المواليد والوفيات ، ثم تلي عمل جرونت بكثير من الدراسات حول الوفيات ونسب ومعدلات الأمراض وأحجام السكان والدخول ونسب البطالة.

تعتمد المجتمعات الكبيرة والحكومات على الدراسات الإحصائية **كموجة أو مرشد** في

عمليات **الدراسة** وأخذ **قرارات** مستقبلية معينة

و

إصدار **توجيهات** خاصة ببعض المشاكل

على سبيل المثال

تقدير حجم البطالة وتقدير حجم التضخم ومعدلات المواليد والوفيات وأسباب الضعف في العملية التعليمية والاقتصادية.

٢-١ تعريف علم الإحصاء

يمكن تعريف علم الإحصاء

علي انه العلم الذي يبحث في

الأساليب المختلفة

لجمع وعرض

و تبويب و تحليل

البيانات

الموجودة داخل

العينة

المأخوذة

من

المجتمع

موضع

الدراسة

حتى يمكن

فهمها و استخدامها في

الوصول الى

نتائج

وصناعة

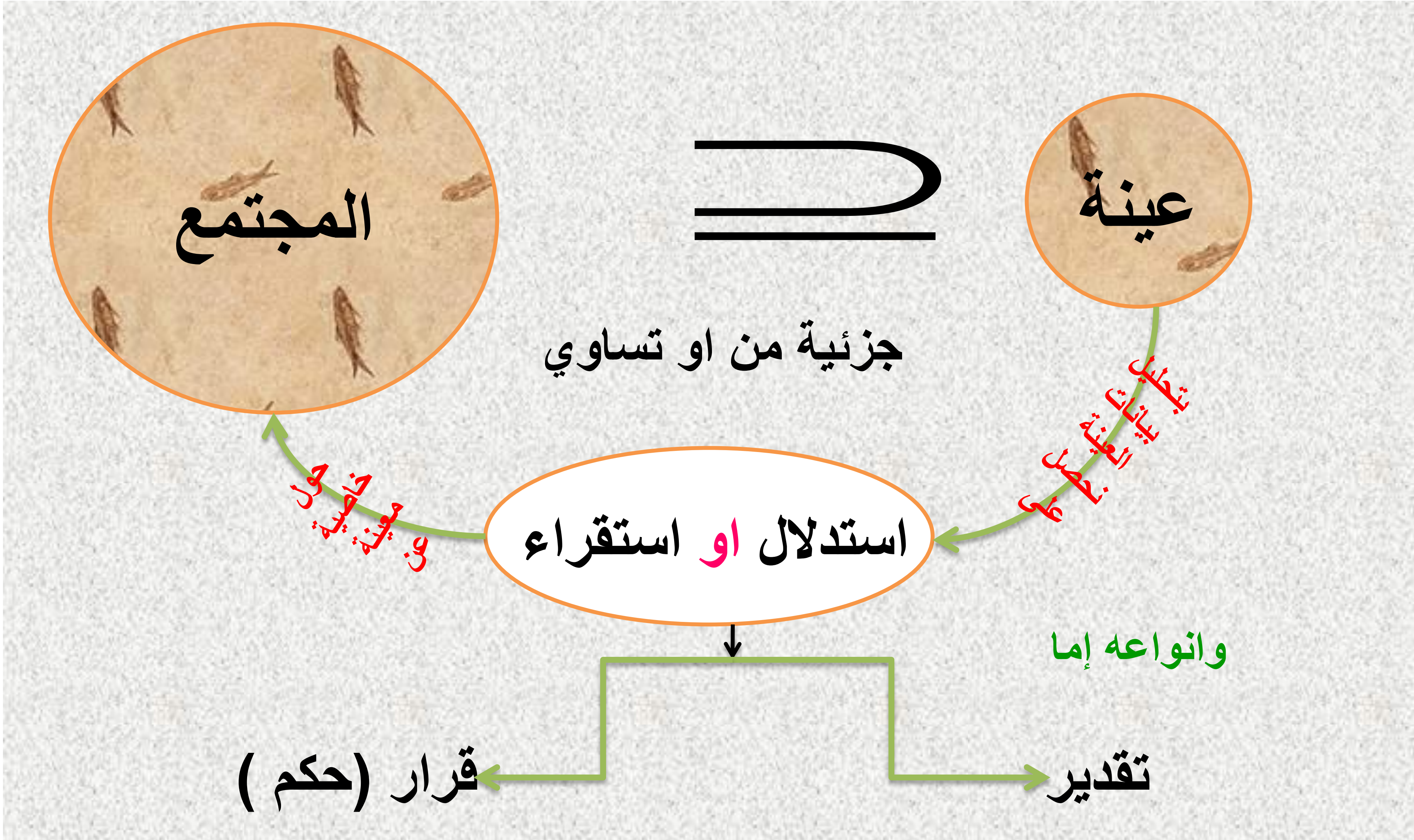
القرارات

حول خاصية معينة

موجودة


المأخوذ منه


والشكل التالي يوضح مفهوم
الإحصاء الاستدلالي


عينة
⊇
المجتمع
جزئية من او تساوي
بتحليل بيانات العينة نحصل على
استدلال او استقراء
حول خاصية معينة عن
وانواعه إما
تقدير
قرار (حكم)

وعليه يمكن تعريف كل من :

١-الإحصاء الوصفي

على أنه

مجموعة الطرق والأساليب التي تستخدم في تنظيم وتلخيص (عرض و تبويب) البيانات.

٢-الإحصاء الاستدلالي

على أنه

مجموعة الطرق والأساليب التي تستخدم في تعميم نتائج العينة على المجتمع التي سحبت منه.

١-٣ الهدف من علم الإحصاء

١- يحتوى علم الإحصاء على

<u>عملية تجميع البيانات والمعلومات</u>

من

<u>محتوى كبير للبيانات</u>

يسمى

<u>بالمجتمع</u> .

٢ – **المجتمع** هو

المجموعة الكاملة من العناصر موضع الدراسة

(يقوم الإحصائي او الباحث بتعيينها على حسب الخاصية المراد دراستها).

٣ - **مثال** :

- مجموعة الطلبة في فرقة دراسية معينة.
- مجموعة درجات طلاب الفرقة الأولى.
- مجموعة الطالبات الدارسات لمقرر معين.

٤ – **العينة هي**

مجموعة جزئية من المجتمع موضع الدراسة .

٥ ـ عملية **الاستقراء أو الاستدلال الإحصائى** حول

خاصية معينة في المجتمع

يمكن الحصول عليها من

البيانات والمعلومات الموجودة

داخل

العينة الممثلة للمجتمع .

٦ – الاستقرار أو الاستدلال الإحصائي يحتوى إما على **تقدير** أو **قرار** (أحكام – تعميمات)

حول خاصية معينة عن **المجتمع**.

٧ـ **المعلم** هو **قياس عددي** يوضح او يصف خاصية معينة عن **المجتمع** ويعرف باسم (**إحصاء مجتمع**).

٨ـ **الإحصاءة** هو **قياس عددي** يوضح او يصف خاصية معينة عن **العينة** ويعرف باسم (**إحصاء عينة**).

٩-مثال:

- في تجربة لمعرفة **أثر استخدام** الحاسب الألى وتطبيقاته على عملية **تدريس مناهج اللغة الانجليزية في المرحلة الابتدائية** .

أجريت بعض الاختبارات على **طلاب عشرة مدارس** في هذه المرحلة وتم تسجيل **نتائج هذه الاختبارات**.

- **القياسات (نتائج الاختبارات)** التي تم تسجيلها عن **طلاب المدارس العشرة** تمثل **عينة** أخذت من **مجتمع الدراسة** (***المدارس التي تستخدم الحاسب الألى في عملية تدريس اللغة الانجليزية في المرحلة الابتدائية***).

- **البيانات الموجودة في هذه العينة** يمكن استخدامها لعمل **استقراء أو استدلال حول** **خاصية معينة** في **المجتمع** ***موضع الدراسة*** .

ومما سبق نستنتج ان :

الهدف من علم الإحصاء الاستدلالي

هو **صناعة الاستدلال او الاستقراء** حول **المجتمع** مستخدماً **البيانات** الموجودة داخل **العينة المأخوذة من هذا المجتمع**.

تمت

بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ ـ ١٤٣١هـ

المحاضرة الثانية

عناصر المحاضرة

2-1 عناصر المشكلة الإحصائية.

2-2 طبيعة القياسات.

2-3 طرق اخذ العينات .

تذكر

علمنا من البند السابق أن الهدف من علم الإحصاء هو عملية الاستدلال حول خاصية معينة عن المجتمع وهذه العملية تعتمد على ثلاثة عناصر سنسميها <u>بعناصر المشكلة الإحصائية</u>

(<u>خطوات حل المشكلة الإحصائية</u>)

٢ـ١ عناصر المشكلة الإحصائية

خطوات حل المشكلة الإحصائية (منهجية علم الإحصاء)

الخطوة الاولى
إختيار العينة (المعاينة)
(جمع البيانات)

الخطوة الثانية
- تنظيم وعرض وتلخيص البيانات
- تحليل البيانات

الخطوة الثالثة
صناعة الاستدلال (استخلاص القرار)

الخطوة الاولى لحل المشكلة الإحصائية

هي عملية الدراسة حول أكثر الطرق اقتصادا

للحصول على كمية معينة من البيانات والمعلومات وتسمى هذه الطريقة (بعملية اختيار العينة أو المعاينة) أو (تخطيط أو تصميم التجربة)

تكلفة الحصول على الكمية المطلوبة من البيانات تتغير كثيرا تبعا للطريقة المتبعة لجمع هذه البيانات

- من أهم وسائل جمع البيانات :

(الاستمارات البحثية – أسلوب الملاحظة – ....).

- من أهم مصادر جمع البيانات :

- المصادر الميدانية (ميدان الظاهرة موضع الدراسة)
- المصادر التاريخية (هيئة البيانات – مؤلفات ودوريات علمية - ...)

الخطوة الثانية لحل المشكلة الإحصائية

تحتوى هذه الخطوة على:

-عملية تنظيم وعرض وتلخيص البيانات .

-عملية تحليل و تعميم البيانات (الحصول على أكبر كم من المعلومات من خلال بيانات العينة).

الخطوة الثالثة لحل المشكلة الإحصائية

تحتوى على عملية استخدام نتائج التحليل السابق لعمل

(استقراء او استدلال) إحصائي عن المجتمع موضع الدراسة (استخلاص القرار).

2-2 طبيعة القياسات

تصنف القياسات (البيانات) إلى

- القياسات النوعية
  (الطبقية أو الوصفية)
- القياسات الكمية
  - قياسات كمية منفصلة
  - قياسات كمية متصلة

القياسات النوعية (أو الطبقية أو الوصفية)

هى التي يمكن تقسيمها أو فصلها إلى طبقات أو أنواع من القياسات يمكن التمييز بينها ببعض الخصائص غيرالعددية مثل:

-مستوى الدراسة -الجنس (ذكر – انثى)

-نوع المنتج من شركة ما للإلكترونيات (راديو- تلفزيون -....)

القياسات الكمية

تحتوى على مجموعة من الأرقام تمثل مقدار القياس أو الخاصية لشيْ معين مقارنة بمجموعته مثل:

عدد أجهزة الراديو التي تنتجها شركة ما للإلكترونيات في أشهر مختلفة.

وتصنف القياسات الكمية إلى:

قياسات كمية منفصلة

هى التي تأخذ قيما محددة مثل:

- عدد المواد الدراسية في الفرقة الأولى.
- عدد أفراد الأسرة.
- ترتيب أخ بين أخوانه وهكذا ......

قياسات كمية متصلة

هى التي تأخذ أي قيم داخل مدى معين أو داخل فترة معينة مثل:

- العمر -الطول - الوزن - درجة الحرارة

وهكذا....

وهناك طريقة أخرى لتصنيف القياسات وهى استخدام المستويات الأربعة للقياس:

- المستوى الاسمي (التصنيفي) للقياس.
- المستوى الرتبي (الترتيبي – التفضيلي) للقياس.
- المستوى الفئوي (الفترى - الفترة) للقياس.
- المستوى النسبي (النسبة) للقياس.

المستوى الاسمي (التصنيفي) للقياس

يتميز هذا النوع بالقياسات التي تحتوى على **الأسماء** ، **العناوين** ، أو **الأصناف** فقط.

وفي هذا المستوى **لا يمكن ترتيب القياسات** بأي طريقة.

امثلة

١- **تصنيف الأفلام** على حسب نوعها (**كوميدي** – **رومانسي** – **تاريخي**).

٢- **تصنيف الأعضاء** الذين تم حضورهم الاقتراع للتصويت على موضوع ما (45 عضوا **ديمقراطيا** – 80 عضوا **جمهوريا** – 90 عضوا **مستقلا**)

٣- **التصنيف** حسب النوع (**ذكور** – **إناث**)

المستوى الرتبي (الترتيبي - التفضيلي) للقياس

يتميز هذا النوع بأنه يحتوى على القياسات التي يمكن إجراء عمليات الترتيب عليها ،

ولكن الفروق بين القياسات (الرتب) إما أن تكون لايمكن تعيينها (تحديد قيمتها) أو تكون لامعنى لها.

امثلة

١- في عينة من منتج معين حجمها 36 ، ثم تصنيف "12" منتجا بحالة جيدة، "16" منتجا بحالة متوسطة ، "8" منتجات بحالة سيئة.

٢- تبعا للتقارير عن جودة أداء المهام في عمل معين تم تصنيف العامل س في المرتبة الثالثة ، العامل ص في المرتبة السابعة ، والعامل ع في المرتبة العاشرة .

المستوى الفئوي (الفترة - الفتري) للقياس

يشبه هذا المستوى إلى حد كبير المستوى الرتبى مع تمييزه بخاصية إضافية وهى

إمكانية تحديد الفروق بين القياسات ومعرفة دلالتها

ولا يوجد لهذا المقياس صفرا أو (نقطة بداية) محددة بل تكون دائما اختيارية أو افتراضية.

مثال

١-الأجسام التي درجة حرارتها ، ٩٨.٢ و ٩٨.٦ درجة فهرنهايت . تتبع المستوى الفئوي للقياس، ونلاحظ أن :

- هذه القيم يمكن ترتيبها،

- يمكن تحديد الفروق بينها،

-عدم وجود صفر مطلق أو نقطة بداية طبيعية لهذه القياسات.

فالقيمة صفر فهرنهايت تبدو كنقطة بداية لكنها اختيارية أو افتراضية وكذلك

لا تعنى عدم وجود حرارة أي لا تعنى غياب الخاصية.

من الخطأ أن نقول أن درجة الحرارة ٥٠ فهرنهايت هي ضعف درجة الحرارة ٢٥ فهر نهايت.

المستوى النسبي (النسبة) للقياس

يعتبر هذا المستوى تطويرا للمستوى الفئوي حيث إنه يحتوى على

نقطة بداية طبيعية (الصفر المطلق)

والذي يعنى غياب الخاصية ، كما أن

الفروق والنسب بين القياسات

في هذا المستوى لها دلالة ومعنى.

امثلة

١- أوزان البلاستيك المهملة من بعض الشركات .

٢- المسافات التي تقطعها السيارات في اختبار لاستهلاك الوقود .

نلاحظ أن القيم في هذه الأمثلة :

- يمكن ترتيبها. - يمكن حساب الفروق بينها.

- وجود صفر مطلقا (نقطة بداية طبيعية).

والذي يعطى **معنى للنسبة** بين القياسات ، فقيمة الوزن 200 كجم هي **ضعف** قيمة الوزن 100 كجم بينما درجة الحرارة ٥٠ فهرنهايت ليست ضعف درجة الحرارة ٢٥ فهرنهايت كما ذكرنا سابقا.

٢-٣ طرق اخذ العينات – انواع العينات

عملية اخذ العينة (المعاينة) عادة تتطلب الكثير من

الوقت | و الجهد | و المال

اكثر من عملية تحليل بيانات هذه العينة.
ويمكن تقليل هذه الأشياء بعملية

التخطيط الجيد والحذر لعملية المعاينة.

انواع العينات

عينات احتمالية

- العينات العشوائية
- العينات الطبقيـــة
- العينات المنتظمة
- العينات العنقودية
- العينات الميسرة

عينات غير احتمالية

- العينة العرضية
- العينة الهادفة (القصدية)

اولآ:العينات احتمالية

١-العينة العشوائية

وفيها **يتساوى** جميع أفراد المجتمع في فرصة اختيارهم داخل العينة.

تعتبر العينة العشوائية عينة ممثلة لمجتمع الدراسة في جميع خصائصه ونادرا ما يحدث اختلاف بين (إحصائيات المجتمع) و (إحصائيات العينة)

مثال : استخدام الحاسب الألى في عملية تخليق أرقام التليفونات.

٢- العينة الطبقية

ويتم في هذا النوع

- تقسيم المجتمع إلى مجتمعين جزئيين (طبقيين) على الأقل بشرط اشتراكهم فى نفس الخاصية، ثم يتم
- اخذ العينة من كل مجتمع جزئي (طبقة).
- من الخصائص التي يمكن تقسيم المجتمع عن طريقها إلى طبقات (مجتمعات جزئية) ، الجنس (ذكر– أنثى)، المستوى الدراسي (ابتدائي – إعدادي – ثانوي- .....)، الحالة الاجتماعية (متزوج – أعزب) ، الحالة الاقتصادية ، .... الخ.

٣- العينة المنتظمة

يتم في هذا النوع

اختيار **نقطة بداية**

ثم اختيار عناصر العينة **<u>بشكل دوري</u> بداية من هذه النقطة**

كأن نختار كل **خامس** طالب إذا كان طول الدورة خمسة أو **ثامن** كتاب إذا كان طول الدورة ثمانية، وهكذا.

٤- العينة العنقودية

يتم في هذا النوع

- **تقسيم المجتمع** إلى **قطاعات** أو **أقسام** تسمى **عناقيد** (*ليست ذات خصائص* **مشتركة** *كما في حالة العينة الطبقية*) ثم يتم

- الاختيار **عشوائيا لعدد من هذه العناقيد** ثم اختيار **كل أعضاء هذه العناقيد في العينة**.

امثلة

١ـ الكليات تشكل عناقيد.

٢ـ الفرق من نفس المستوى في الكلية الواحدة تشكل عناقيد.

٣ـ سكان الأحياء المختلفة داخل مدينة أو قرية واحدة يشكلون عناقيد.

٥- العينة الميسرة (الجاهزة)

وفيها يتم استخدام النتائج المتاحة من قبل بسهولة و بسرعة.

في بعض الحالات تكون العينات الجاهزة جيدة وفى بعض الحالات تحتوى على انحرافات (لصغر حجم العينة).

امثلة

١- استخدام إحدى الشركات للبيانات سابقة التجهيز من قبل إحدى الهيئات أو المؤسسات.

٢- استخدام الباحث لقواعد البيانات العلمية في تخصص معين.

٣- الاستعانة بمراكز المعلومات في التخصصات المختلفة.

ثانياً:العينات غير الإحتمالية

وهى العينات التي يتم أخذها تحت شروط ومواصفات أو معايير يراها الباحث لتحقيق غرض معين في التجربة.

وبذلك فإن هذا النوع لا يتبع نظرية الاحتمالات في الاختيار.

هذا النوع من العينات يكون أكثر فائدة من ناحية الحصول علية حيث أنة لا يحتاج إلى الجهد والتكاليف والوقت ، كما في الأنواع السابقة ، إلا إنه من الصعب تعميم نتائجه على المجتمع.

١- العينة العرضية

ويتم اخذ مثل هذه العينات عن **طريق الصدفة** . **ولا يمكن تعميم** نتائجها على مجتمع الدراسة .

مثال ذلك ، أن يقوم الباحث بتوزيع **استبيان** خاص به على العاملين أثناء مشاهدتهم لمباراة كرة القدم أو أثناء تناولهم وجبة الإفطار.

٢ـ العينة الهادفة (القصدية)

وفى هذا النوع يقوم الباحث باختيار العينة تحت شروط معينة لتحقيق الهدف أو الغرض من التجربة أو الدراسة.

ومثال ذلك ، إذا أراد باحث الكتابة عن حدث معين لم يعاصره كقيام احد الثورات أو استقلال احد الشعوب فإنه لابد من اختيار أفراد عينته من أشخاص قد عاصروا هذه الفترة وعلى قدر من الوعي والموضوعية ليحصل منهم على البيانات أو المعلومات التي يراها.

أخطاء جمع البيانات (أخطاء المعاينة)

تتعرض البيانات لنوعين من الأخطاء عند تجميعها هما:

خطأ التحيز (*Error of Bias*):

مصدره : الباحث نفسه أو مفردات المجتمع موضع الدراسة .

امكانية حدوثه : عند استخدام اسلوب الحصر الشامل أو العينة العشوائية .

خطأ المعاينة العشوائية (*Error of Bias Sampling*):

مصدره : الصدفة (ليس الباحث أو افراد المجتمع (المبحوث)) .

امكانية حدوثه : عند استخدام اسلوب العينة العشوائية .

تمت

بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ ـ ١٤٣١هـ

المحاضرة الثالثة

عناصر المحاضرة

الطرق البيانية لوصف القياسات (البيانات)

١-١ الهدف من وصف القياسات (البيانات)
١-٢ طرق الوصف
١-٣ الطريقة البيانية (الغرض من استخدامها)
١-٤ الجداول التكرارية (انواعها وطرق التأكد من عملية الجدولة)

نحتاج في هذه المحاضرة

رمز التجميع $\sum_{i=1}^{k} f_i$

إشارة لعملية **جمع** عدد k من الحدود بداية من الحد رقم (١) الى الحد رقم (k)

$$\sum_{i=1}^{k} f_i = f_1 + f_2 + \dots\dots\dots + f_k$$

الحد الاول (رقم ١) بوضع $i=1$

الحد الاخير (رقم k) بوضع $i=k$

1-1 الهدف من وصف القياسات (البيانات)

لماذا نصف مجموعات القياسات ؟

- بالرجوع إلى الهدف من علم الإحصاء ، صناعة الاستدلال الإحصائي أو الاستقراء الإحصائي ، حول خاصية معينة في مجتمع الدراسة مستخدمين البيانات الموجودة بالعينة المختارة.

- لكى نقوم بعمل جملة استدلالية فإننا نحتاج إلى طريقة لوصف هذا المجتمع.

طرق الوصف

- الطريقة البيانية (**الرسوم**)
- الطريقة العددية (**الحسابية**)

عند تحليل مجموعة كبيرة من القياسات ، فإننا نقوم أولا **بترتيب** و**تلخيص** هذه البيانات مستخدمين **الجداول و الأشكال البيانية**.

الطرق البيانية

تستخدم الطريقة البيانية لعرض مجموعة القياسات في صورة أشكال أو رسومات بيانية تعطى الدارس أو الباحث وصفا مرئيا كافيا عن هذه المجموعة من القياسات.

الجدول التكراري

يستخدم الجدول التكراري لسرد أو تلخيص القياسات داخل فترات (أو فصول) تبعا لعدد مرات (أو تكرار) وقوع البيان داخل كل فترة (أو فصل).

الحد الأدنى للفترة (أو الفصل) هو **أصغر** عدد تحتويه الفترة (أوالفصل).

الحد الأعلى للفترة (أو الفصل) هو **اكبر** عدد تحتويه الفترة (أو الفصل).

اولآ... سنعرض الاشكال المختلفة للجدول التكراري

شكل ١ – الجدول التكراري

رقم الفترة (الفصل) i	حدود الفترة	التكرارات المطلقة: علامات التفريغ	التكرارات المطلقة: التكرار f_i
١	الحد الادني - الحد الاعلى	I	f_1 =١
٢	الحد الادني - الحد الاعلى	II	f_2 =٢
٣	الحد الادني - الحد الاعلى	III	f_3 =٣
....	....	....	....
المجموع	....	....	$\sum_{i=1}^{k} f_i = n$

f_i تكرار الفترة (الفصل) رقم i أي ان f_1 = تكرار الفترة (الفصل) رقم 1

وهكذا........

يمكن التأكد من صحة عملية الجدولة في الجدول التكراري السابق بأخذ المجموع الكلي لعمود التكرارات نجد انه يساوي عدد البيانات

اي ان

$$\sum_{i=1}^{k} f_i = f_1 + f_2 + \dots\dots\dots + f_k = n$$

حيث n يمثل العدد الكلي للبيانات و k يمثل عدد الفترات

نموزج اخر من الجدول التكراري يمكن الحصول عليه باستبدال التكرارات المطلقة بالتكرارات النسبية.

التكرار النسبي للفترة رقم i يعطي بالمعادلة $\tilde{f}_i = \frac{f_i}{n}$

حيث f_i تمثل تكرار الفترة رقم i

و n تمثل العدد الكلي للبيانات.

(شكل 2 – الجدول التكراري النسبي)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات النسبية $\tilde{f}_i = \frac{f_i}{n}$
١	الحد الادني - الحد الاعلى	التكرار المطلق للفترة / العدد الكلي للبيانات (n)
٢	الحد الادني - الحد الاعلى	التكرار المطلق للفترة / العدد الكلي للبيانات (n)
٣	الحد الادني - الحد الاعلى	التكرار المطلق للفترة / العدد الكلي للبيانات (n)
....	....	....
....	....	....
المجموع	....	$\sum_{i=1}^{k} \frac{f_i}{n} = 1$

وبالمثل يمكن التأكد من صحة عملية الجدولة في الجدول التكراري النسبي بأخذ المجموع الكلي لعمود التكرارات النسبية نجد انه يساوي الواحد الصحيح

اي ان

$$\sum_{i=1}^{k}\frac{f_i}{n}=\frac{f_1}{n}+\frac{f_2}{n}+.........+\frac{f_k}{n}=$$

$$=\frac{f_1+f_2+.........+f_k}{n}=\frac{n}{n}=1$$

حيث n يمثل العدد الكلي للبيانات و k يمثل عدد الفترات

نموزج ثالث من الجدول التكراري يمكن الحصول عليه باستبدال التكرارات المطلقة بالتكرارات التراكمية (التجميعية) .

التكرار التراكمي للفترة i يساوي مجموع تكرارات هذه الفترة وما قبلها (ما يسبقها) من فترات.

لاحظ ان التكرار التراكمي للفترة الاولى هو نفس التكرار المطلق لهذه الفترة.

(شكل ٣ – الجدول التكراري التراكمي (التجميعي))

رقم الفترة (الفصل) i	حدود الفترة	التكرارات المطلقة	التكرارات التراكمية (التجميعية)
١	الحد الادني - الحد الاعلى	١	١
٢	الحد الادني - الحد الاعلى	٢	٣
٣	الحد الادني - الحد الاعلى	٣	٦
....	....	....	....
k	....	....	n

وبالمثل يمكن الحصول علي التكرار التراكمي النسبي كما هو موضح بالجدول التالي (شكل ٤)

وبالمثل يمكن التأكد من صحة عملية الجدولة في الجدول التكراري التراكمي عن طريق التأكد من وجود العدد الكلي للبيانات n امام الفترة الاخيرة (الفترة k) في عمود التكرارات التراكمية.

(شكل 4 – الجدول التكراري التراكمي النسبي)

التكرارات التراكمية (النسبية)	التكرارات التراكمية	التكرارات المطلقة	حدود الفترة	رقم الفترة (الفصل) i
التكرار التراكمي للفترة / العدد الكلي للبيانات(n)	١	١	الحد الادني - الحد الاعلى	١
التكرار التراكمي للفترة / العدد الكلي للبيانات(n)	٣	٢	الحد الادني - الحد الاعلى	٢
التكرار التراكمي للفترة / العدد الكلي للبيانات(n)	٦	٣	الحد الادني - الحد الاعلى	٣
....	....	....	....	....
....	....	....	....	....
1	....	....	....	k

وبنفس الطريقة يمكن التأكد من صحة عملية الجدولة في الجدول التكراري التراكمي النسبي عن طريق التأكد من وجود العدد واحد امام الفترة الاخيرة (الفترة *k*) في عمود التكرارات التراكمية النسبية.

تمت
بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ – ١٤٣١هـ

المحاضرة الرابعة

عناصر المحاضرة

تابع الطرق البيانية لوصف القياسات (البيانات)

١ ـ ١ خطوات الطريقة البيانية
٢ـ ١ المدرجات التكرارية وشكل القياسات

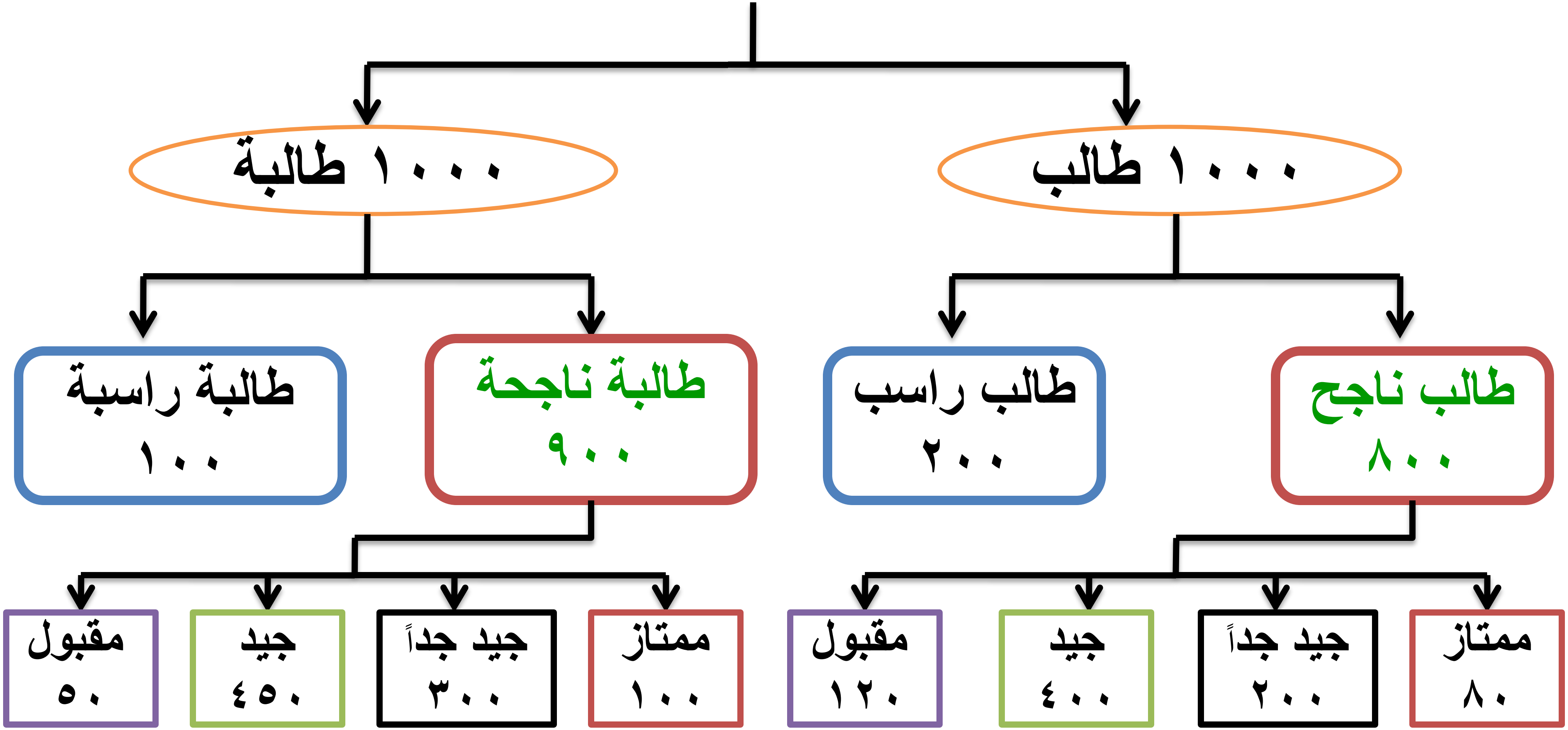
تمهيد

بفرض مجتمع الدارسين بالفرقة الاولي بكلية الاداب

بفرض ان عددهم ٢٠٠٠ دارس

لدراسة هذا المجتمع تبعاً لنتيجة نهاية الفصل الدراسي الاول

مجتمع الدارسين (٢٠٠٠)



وتكون الدراسة اكثر وضوحاً عندما نضيف الي وصف المجتمع المعلومة التالية :

طالب ناجح
٨٠٠

- **ممتاز ٨٠**
  - من ٩٥- ١٠٠ (٢٠)
  - من ٩٠- ٩٥ (٦٠)
- **جيد جداً ٢٠٠**
  - من ٨٥- ٩٠ (٨٠)
  - من ٨٠- ٨٥ (١٢٠)
- **جيد ٤٠٠**
  - من ٧٥- ٨٠ (١٨٠)
  - من ٧٠- ٧٥ (٢٢٠)
- **مقبول ١٢٠**
  - من ٦٥ – ٧٠ (٥٥)
  - من ٦٠ – ٦٥ (٦٥)

طالبة ناجحة
٩٠٠

- **ممتاز ١٠٠**
  - من ٩٥- ١٠٠ (٤٥)
  - من ٩٠- ٩٥ (٥٥)
- **جيد جداً ٣٠٠**
  - من ٨٥- ٩٠ (١٨٠)
  - من ٨٠- ٨٥ (١٢٠)
- **جيد ٤٥٠**
  - من ٧٥- ٨٠ (٢٤٠)
  - من ٧٠- ٧٥ (٢١٠)
- **مقبول ٥٠**
  - من ٦٥ – ٧٠ (٣٥)
  - من ٦٠ – ٦٥ (١٥)

١-١ خطوات الطريقة البيانية

سنقوم بشرح مفهوم الطريقة البيانية من خلال هذا المثال:

مثال

في اختبار لقياس درجة الذكاء يحتوى على 30 فقرة رصدت درجات 25 طالبا على النحو التالي:

(لاحظ ان حجم البيانات = ٢٥ (عدد الطلاب))

درجات الطلاب

25	29	23	27	25
23	22	25	22	28
28	24	17	24	30
19	17	23	21	24
15	20	26	19	23

١- نوجد اكبر درجة وهى 30 وكذلك اقل درجة وهى ١٥ واللتان تعطيان مؤشرا إلى أن مدى هذه القياسات هو 15 (الفرق بين اكبر درجة و اقل درجة).

٢- لمعرفة كيفية توزيع درجات الطلبة بين الحد الأعلى للدرجات (أكبر درجة = ٣٠) والحد الأدنى للدرجات (اقل درجة = ١٥) سنقوم بتقسيم هذه الفترة (المدي) إلى فترات جزئية متساوية أو فصول.

الفترة من 15 إلى 30 يمكن تقسيمها عادة إلى عدد (من 5 إلى 20) فترة جزئية أو فصل حسب عدد القياسات المتاحة.

وعند تحديد عدد الفترات الجزئية أو الفصول يجب مراعاة أن (خطأ التجميع) يزداد بنقصان عدد الفترات الجزئية وبالمقابل فإن زيادة عدد الفترات الجزئية يعنى الاقتراب من التوزيع الأصلى للقياسات.

٣- بفرض إننا نريد الحصول على ٧ فترات جزئية أو فصول ، فإنه لابد من اختيار **طول** مناسب لهذه الفترات الجزئية ، وذلك

بتقسيم مدى هذه القياسات (١٥) على ٧
(عدد الفترات الجزئية المراد الحصول عليها)

٤-**العدد الصحيح** 2 (ناتج عملية القسمة) هو أقرب عدد صحيح يمكن استخدامة كطول للفترة الجزئية محققا شرط اختيار اكبر عدد من الفترات الجزئية أو الفصول.

٥- نحدد النقط الحدية للفترات الجزئية على النحو التالي:

14.5، 16.5، 18.5، 20.5، 22.5، 24.5، 26.5، 28.5 و 30.5

- يجب ان نتأكد من إنه **لا يوجد قياس أو مفردة من البيانات تقع على أي نقطة حدية** ، مما يؤكد أن **كل قياس أو مفردة يقع في فترة جزئية واحدة أو فصل واحد فقط.**

- وتسمى **النقاط المنصفة** لهذه الفترات الجزئية **بمراكز الفترات** أو **علامات الفترات** أي أن :

$$\text{مركز الفترة} = \frac{\text{الحد الادني للفترة} + \text{الحد الاعلي للفترة}}{2}$$

٦- نقوم **بالمسح التدريجي لدرجات الطلاب** كما يلي:

- وضع إشارة (علامة التفريغ) أمام الفترة التي تنتمي إليها كل درجة .

- ثم يتم ترجمة هذه **الإشارات** إلى **أرقام** تعبر عن **تكرار كل فترة (أو فصل)** وتوضع في عمود خاص بها في **الجدول التكراري.**

- كما يمكن ترجمة هذه الإشارات إلى **نسب تكرارية** أو **(احتمالات)** توضع أيضا في عمود خاص بها في الجدول التكراري كما هو موضح **بالجدول** التالي **(شكل ١)**:

(شكل ١ – الجدول التكراري)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات المطلقة: علامات التفريغ	التكرارات المطلقة: التكرار f_i
١	١٤.٥ – ١٦.٥	$\vert$	١
٢	١٦.٥ – ١٨.٥	$\vert\vert$	٢
٣	١٨.٥ – ٢٠.٥	$\vert\vert\vert$	3
٤	٢٠.٥ – ٢٢.٥	$\vert\vert\vert$	3
٥	٢٢.٥ – ٢٤.٥	$\cancel{\vert\vert\vert\vert}\ \vert\vert$	7
٦	٢٤.٥ – ٢٦.٥	$\vert\vert\vert\vert$	4
٧	٢٦.٥ – ٢٨.٥	$\vert\vert\vert$	3
٨	٢٨.٥ – ٣٠.٥	$\vert\vert$	2
المجموع $\sum$	--------	--------	$\sum_{i=1}^{8} f_i = 25$

باستبدال عمود التكرارات المطلقة في الجدول السابق (شكل ١) بعمود التكرارات النسبية نحصل علي الجدول التكراري النسبى (شكل ٢).

(شكل 2 – الجدول التكراري النسبي)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات النسبية $\tilde{f}_i = \frac{f_i}{n}$
١	١٤.٥ – ١٦.٥	١/٢٥
٢	١٦.٥ – ١٨.٥	٢/٢٥
٣	١٨.٥ – ٢٠.٥	٣/٢٥
٤	٢٠.٥ – ٢٢.٥	٣/٢٥
٥	٢٢.٥ – ٢٤.٥	٧/٢٥
٦	٢٤.٥ – ٢٦.٥	٤/٢٥
٧	٢٦.٥ – ٢٨.٥	٣/٢٥
٨	٢٨.٥ – ٣٠.٥	٢/٢٥
المجموع	--------	$\sum_{i=1}^{8} \frac{f_i}{25} = 1$

باستبدال عمود **التكرارات المطلقة** في الجدول (شكل١) بعمود **التكرارات التراكمية** نحصل علي **الجدول التكراري التراكمي** (شكل ٣).

(شكل ٣ – الجدول التكراري التراكمي)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات التراكمية
١	١٤.٥ – ١٦.٥	١
٢	١٦.٥ – ١٨.٥	٣
٣	١٨.٥ – ٢٠.٥	٦
٤	٢٠.٥ – ٢٢.٥	٩
٥	٢٢.٥ – ٢٤.٥	١٦
٦	٢٤.٥ – ٢٦.٥	٢٠
٧	٢٦.٥ – ٢٨.٥	٢٣
٨	٢٨.٥ – ٣٠.٥	٢٥

لاحظ ان قيمة التكرار التراكمي للفترة الاخيرة = ٢٥ (عدد البيانات)

كما يمكننا ايجاد

<u>الجدول التكراري التراكمى النسبى</u> (شكل ٤)

بقسمة التكرار التراكمي لكل فترة علي اجمالي عدد البيانات (٢٥) كما يلي:

(شكل ٤ – الجدول التكراري التراكمي النسبي)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات التراكمية	التكرارات التراكمية النسبية
١	١٤.٥ – ١٦.٥	١	٢٥/١
٢	١٦.٥ – ١٨.٥	٣	٢٥/٣
٣	١٨.٥ – ٢٠.٥	٦	٢٥/٦
٤	٢٠.٥ – ٢٢.٥	٩	٢٥/٩
٥	٢٢.٥ – ٢٤.٥	١٦	٢٥/١٦
٦	٢٤.٥ – ٢٦.٥	٢٠	٢٥/٢٠
٧	٢٦.٥ – ٢٨.٥	٢٣	٢٥/٢٣
٨	٢٨.٥ – ٣٠.٥	٢٥	٢٥/٢٥ = ١

لاحظ ان قيمة التكرار التراكمي النسبي للفترة الاخيرة هو ٢٥/٢٥ = ١

وأخيرا يمكننا ايجاد

الجدول التكراري الكامل (شكل ٥)

وهو الجدول الشامل لجميع الجداول السابقة.

(شكل ٥ – الجدول التكراري الكامل)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات المطلقة f_i	التكرارات النسبية $\tilde{f}_i$	التكرارات التراكمية	التكرارات التراكمية النسبية
١	١٤.٥ – ١٦.٥	١	٢٥/١	١	٢٥/١
٢	١٦.٥ – ١٨.٥	٢	٢٥/٢	٣	٢٥/٣
٣	١٨.٥ – ٢٠.٥	٣	٢٥/٣	٦	٢٥/٦
٤	٢٠.٥ – ٢٢.٥	٣	٢٥/٣	٩	٢٥/٩
٥	٢٢.٥ – ٢٤.٥	٧	٢٥/٧	١٦	٢٥/١٦
٦	٢٤.٥ – ٢٦.٥	٤	٢٥/٤	٢٠	٢٥/٢٠
٧	٢٦.٥ – ٢٨.٥	٣	٢٥/٣	٢٣	٢٥/٢٣
٨	٢٨.٥ – ٣٠.٥	٢	٢٥/٢	٢٥	٢٥/٢٥ = ١
المجموع	-------	٢٥	١	-------	-------

٢-١ المدرجات التكرارية وشكل القياسات

بعد إتمام عملية جدولة البيانات ، يمكننا وصف البيانات باستخدام

المدرج التكراري (رسم التكرار ضد الفترات الجزئية)

أو

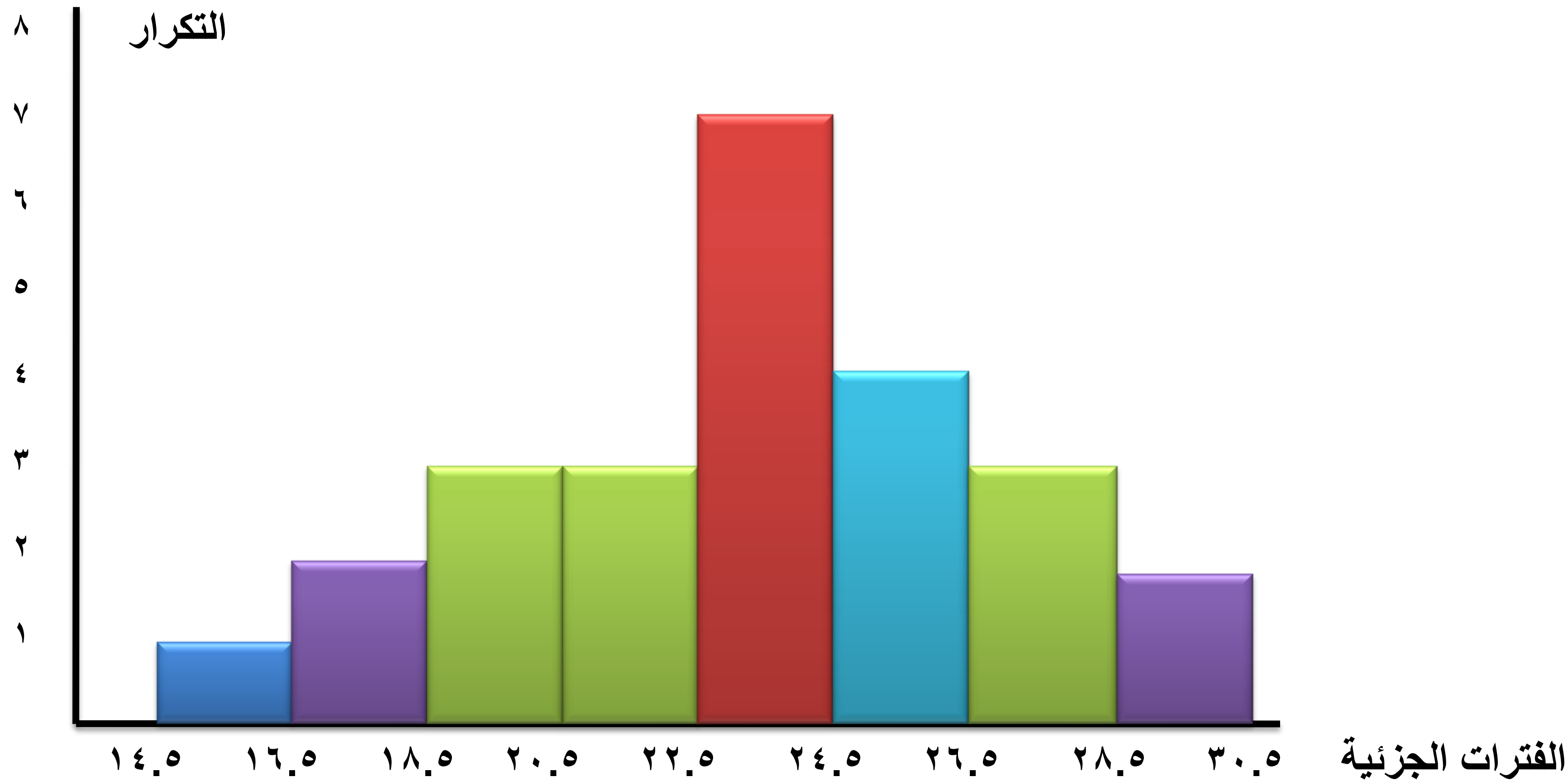
المدرج التكراري النسبي (رسم التكرار النسبي ضد الفترات الجزئية)

باستخدام البيانات في المثال السابق لرسم المدرج التكراري والمدرج التكراري النسبي

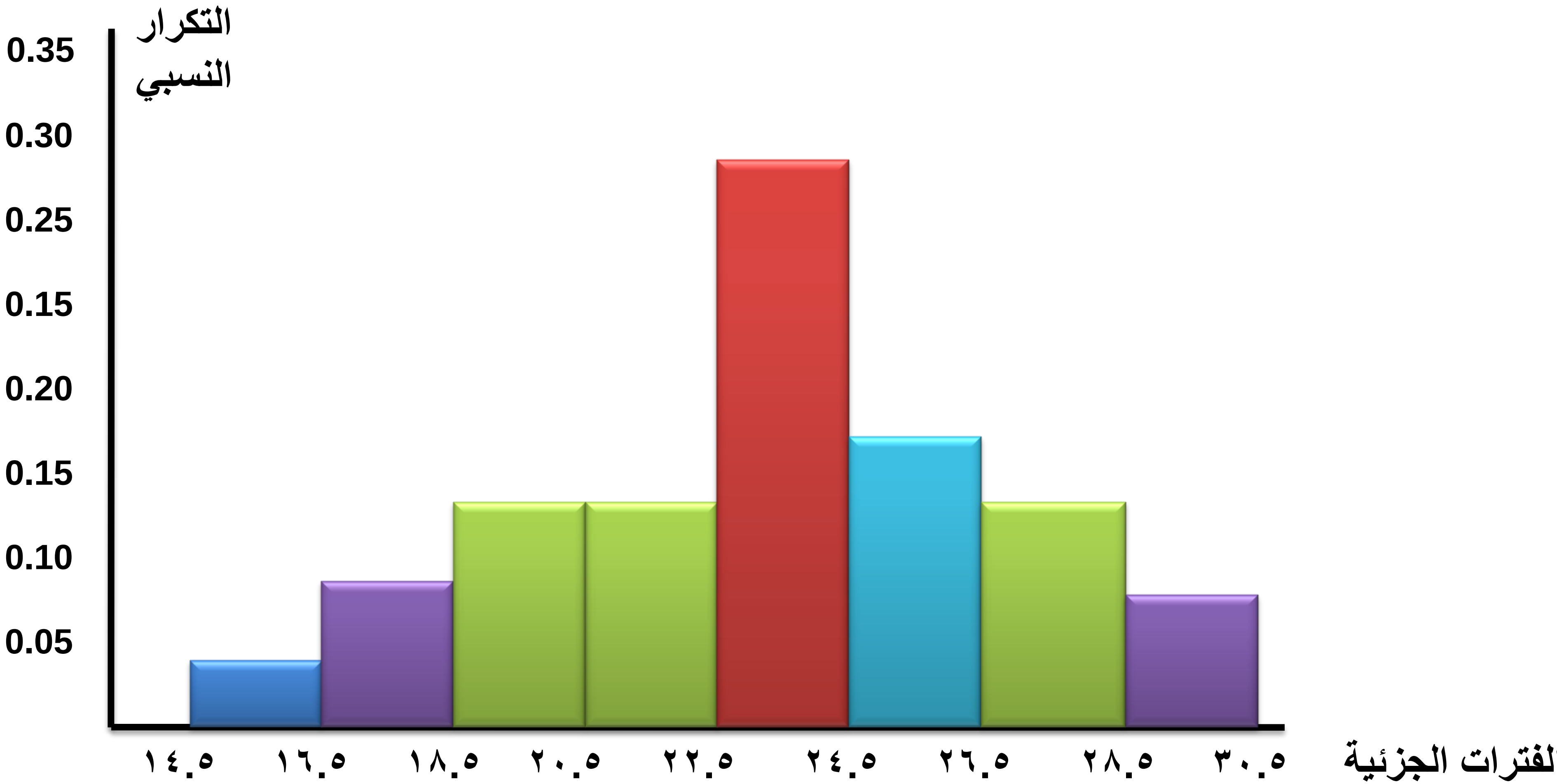
ماذا نلاحظ من الرسم ؟

المدرج التكراري

شكل (١)



شكل (٢) المدرج التكراري النسبي



نلاحظ ان

المدرجين التكراريين **متطابقان** تماماً

اي لهما نفس :

- المحور الافقي
- المستطيلات أعلي الفترات الجزئية

لكنهما يختلفان في

- المحور الرأسي

-المدرج التكراري النسبي عادة يعرف باسم التوزيع التكراري لأنة يعرض طريقة توزيع البيانات) على المحور الأفقي .

-المستطيلات الموجودة أعلى الفترات الجزئية (أو الفصول) في المدرج التكراري النسبي تشير إلى :

- المستطيل أعلى الفترة الجزئية رقم *i* يمثل نسبة البيانات التي تقع داخل هذه الفترة الجزئية .

- المستطيل أعلى الفترة الجزئية رقم *i* يمثل أيضا احتمال أن القيمة المختارة عشوائيا من العينة تقع داخل هذه الفترة الجزئية .

تمت
بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ – ١٤٣١هـ

المحاضرة الخامسة

عناصر المحاضرة

-مراجعة الطريقة البيانية لوصف القياسات.

مراجعة المحاضرة السابقة

بالرجوع الي المثال الموجود في نهاية المحاضرة السابقة .....

المثال

نتذكر

في اختبار لقياس درجة الذكاء يحتوى على 30 فقرة رصدت درجات 25 طالبا على النحو التالي:

(لاحظ ان حجم البيانات = ٢٥ (عدد الطلاب))

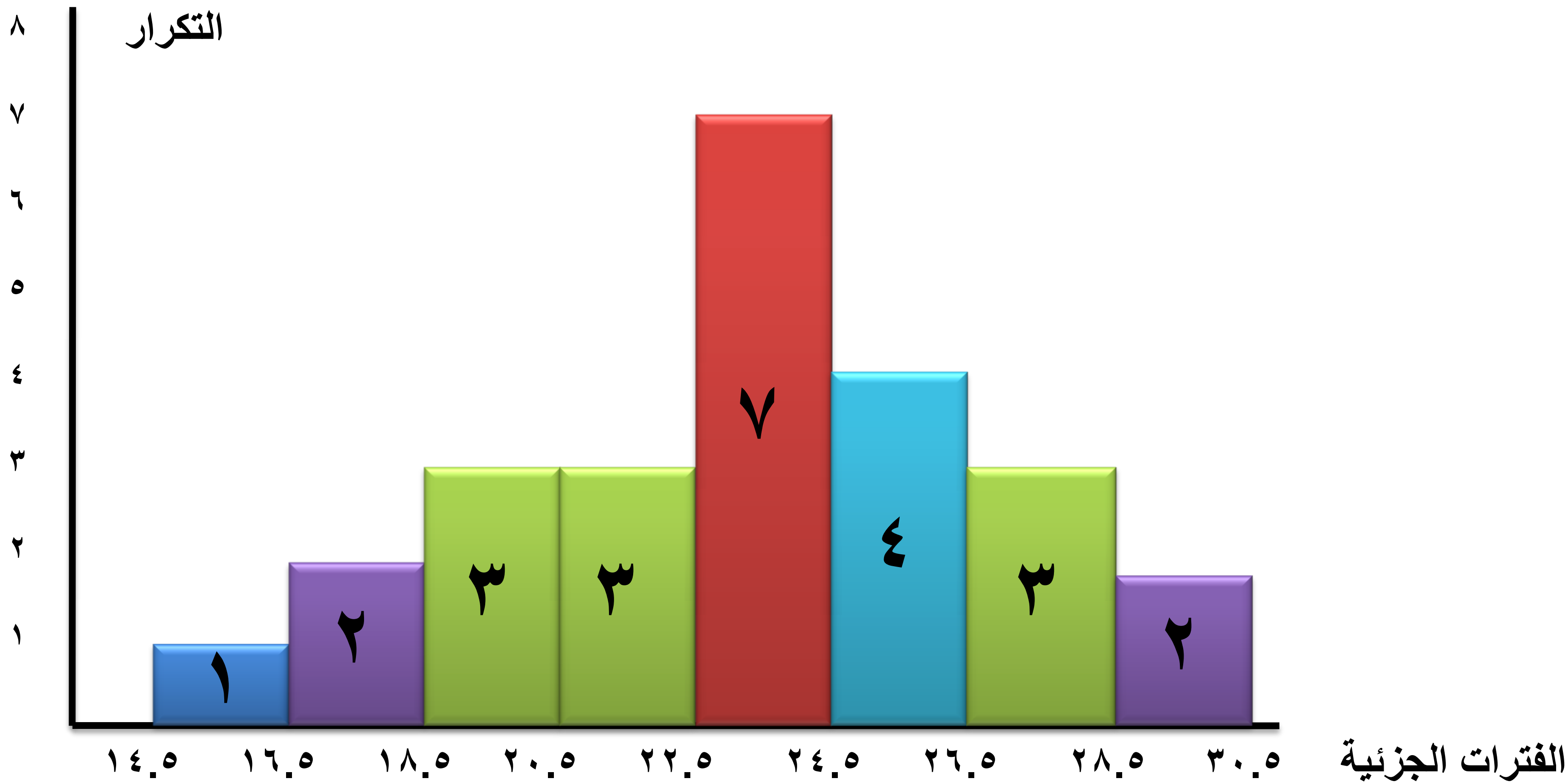
درجات الطلاب

25	29	23	27	25
23	22	25	22	28
28	24	17	24	30
19	17	23	21	24
15	20	26	19	23

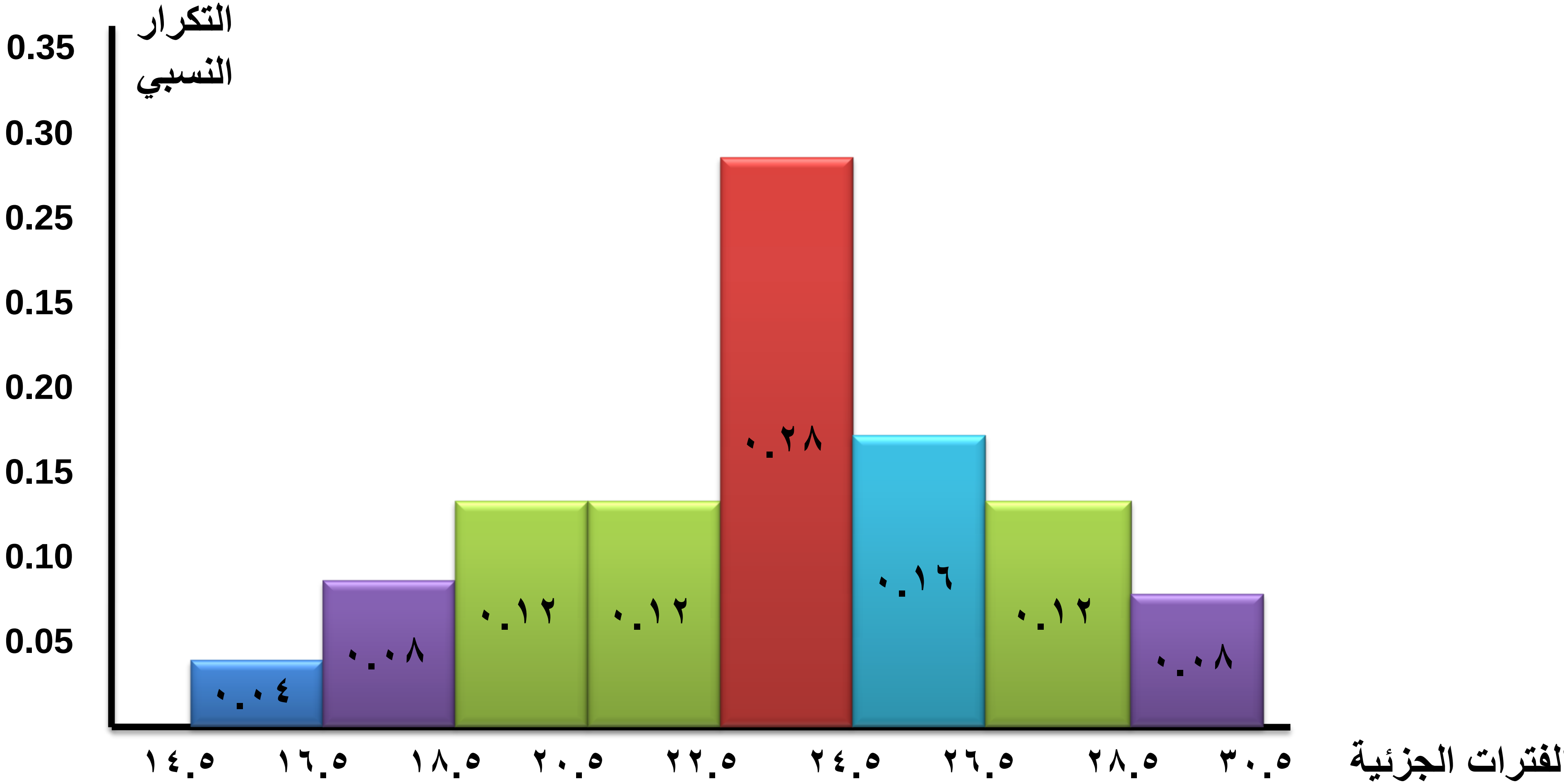
(الجدول التكراري الكامل)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات المطلقة f_i	التكرارات النسبية $\tilde{f}_i$	التكرارات التراكمية	التكرارات التراكمية النسبية
١	١٤.٥ – ١٦.٥	١	٢٥/١= ٠.٠٤	١	٢٥/١= ٠.٠٤
٢	١٦.٥ – ١٨.٥	٢	٢٥/٢= ٠.٠٨	٣	٢٥/٣ = ٠.١٢
٣	١٨.٥ – ٢٠.٥	٣	٢٥/٣ = ٠.١٢	٦	٢٥/٦ = ٠.٢٤
٤	٢٠.٥ – ٢٢.٥	٣	٢٥/٣ = ٠.١٢	٩	٢٥/٩ = ٠.٣٦
٥	٢٢.٥ – ٢٤.٥	٧	٢٥/٧= ٠.٢٨	١٦	٢٥/١٦= ٠.٦٤
٦	٢٤.٥ – ٢٦.٥	٤	٢٥/٤ = ٠.١٦	٢٠	٢٥/٢٠= ٠.٨٠
٧	٢٦.٥ – ٢٨.٥	٣	٥/٣ = ٠.١٢	٢٣	٢٥/٢٣ = ٠.٩٢
٨	٢٨.٥ – ٣٠.٥	٢	٢٥/٢= ٠.٠٨	٢٥	٢٥/٢٥ = ١
المجموع	------	٢٥	١	------	------

المدرج التكراري   شكل (١)

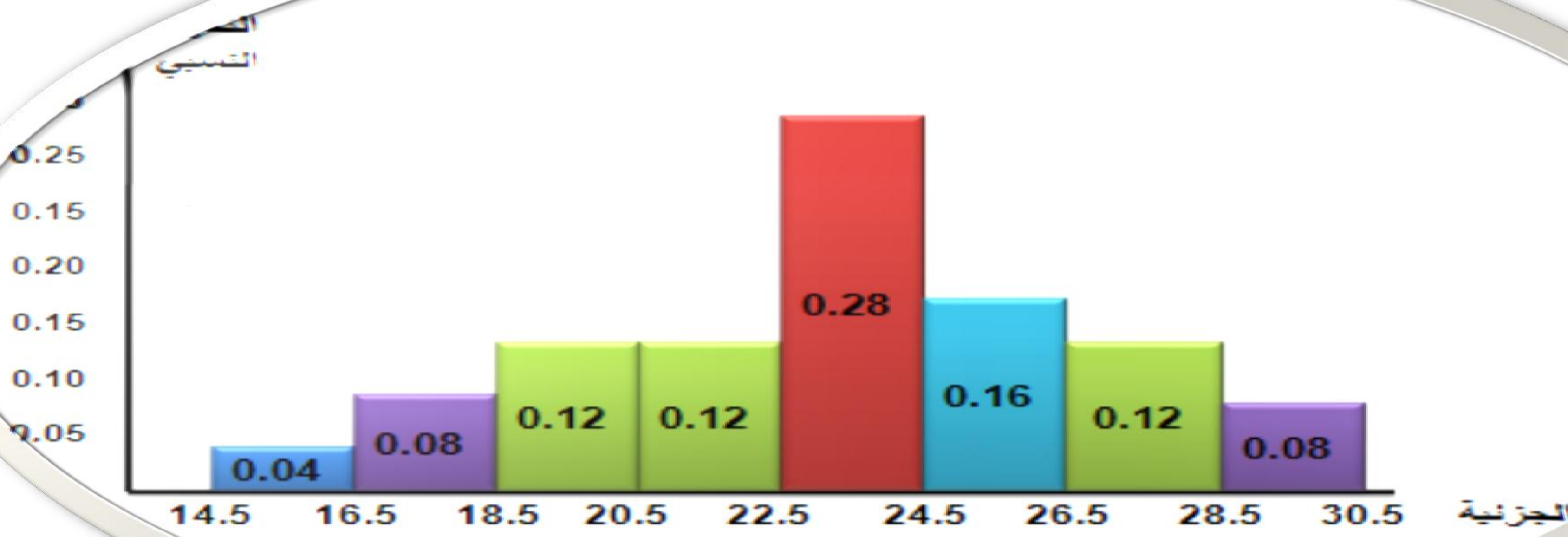


شكل (٢) المدرج التكراري النسبي



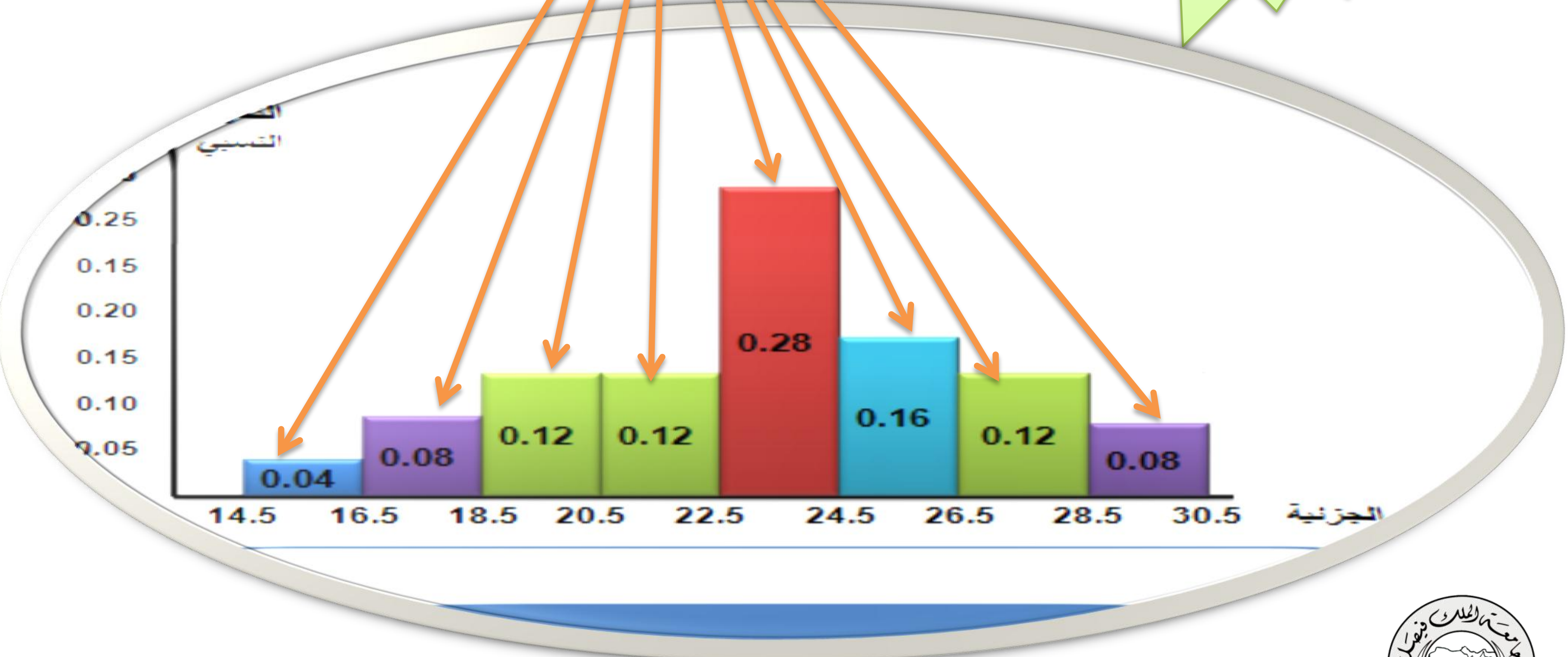
لاحظ

المدرج التكراري النسبي عادة يعرف باسم التوزيع التكراري لأنة (يعرض طريقة توزيع البيانات) على المحور الأفقي .



كل مستطيل يعلو فترة جزئية معينة .. ماذا يمثل؟

لاحظ ايضاً



يمثل أولاً

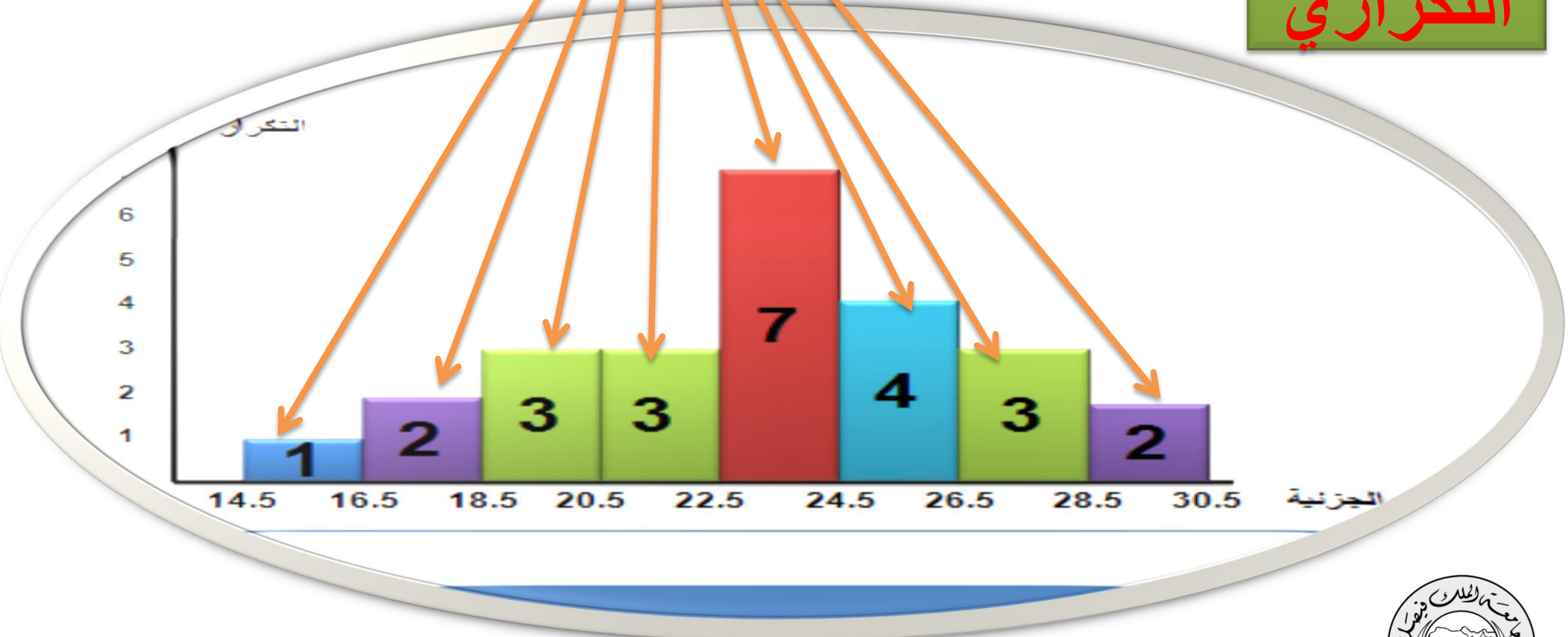
نسبة البيانات التي تقع داخل هذه الفترة الجزئية المعينة.

ثانياً

احتمال أن القيمة المختارة عشوائيا من العينة تقع داخل هذه الفترة الجزئية المعينة.

لكن في المدرج التكراري

كل مستطيل يعلو فترة جزئية معينة ..ماذا يمثل؟



يمثل فقط

<u>عدد</u> البيانات التي تقع داخل هذه الفترة الجزئية المعينة.

مما سبق

نحاول

الاجابة علي الأسئلة الاتية :

١- ما هي نسبة الطلبة الذين حصلوا على درجات أقل من 20.5 ؟

المطلوب هنا

نسبة الطلبة

اذاً نستخدم

الجدول التكراري النسبي

او

المدرج التكراري النسبي

(الجدول التكراري الكامل)

التكرارات التراكمية النسبية	التكرارات التراكمية	التكرارات النسبية $\tilde{f}_i$	التكرارات المطلقة f_i	حدود الفترة	رقم الفترة (الفصل) i
0.04 =25/1	1	0.04 =25/1	1	16.5 – 14.5	1
0.12 = 25/3	3	0.08 =25/2	2	18.5 – 16.5	2
0.24 = 25/6	6	0.12 = 25/3	3	20.5 – 18.5	3

اولاً

من الجدول التكراري النسبي

الثلاث فترات المظللة باللون الاحمر تمثل الطلاب الذين حصلوا على درجات أقل من 20.5.

وبالتالي نسبة عدد الطلاب هي مجموع التكرارات النسبية للفترات الثلاثة (عمود التكرارات النسبية المظلل بالون الاخضر) (0.24 = 0.12+ 0.08 + 0.04).

نلاحظ ان نفس النتيجة يمكن الحصول عليها مباشرة من عمود التكرارات التراكمية النسبية من الخانة الاخيرة (المظللة باللون الاصفر) .

٢ـ ما هي نسبة الطلبة الذين حصلوا على درجات أكبر من 26.5 ؟

(الجدول التكراري الكامل)

رقم الفترة (الفصل) i	حدود الفترة	التكرارات المطلقة f_i	التكرارات النسبية $\tilde{f}_i$	التكرارات التراكمية	التكرارات التراكمية النسبية
7	28.5 – 26.5	3	0.12 = 5/3	23	0.92 = 25/23
8	30.5 – 28.5	2	0.08 =25/2	25	1 = 25/25

بنفس الطريقة السابقة نجد ان العمود المظلل باللون الاحمر يمثل مجموعة الطلاب الذين حصلوا على درجات أكبر من 26.5 .
وبالتالي فإن نسبة الطلبة هي مجموع التكرارات النسبية الممثلة بالعمود المظلل باللون الاخضر
(0.12 + 0.08 = 0.20).

٣ـ ما هي نسبة الطلبة الذين حصلوا على درجات بين 20.5 و 26.5؟

(الجدول التكراري الكامل)					
رقم الفترة (الفصل) i	حدود الفترة	التكرارات المطلقة f_i	التكرارات النسبية $\tilde{f}_i$	التكرارات التراكمية	التكرارات التراكمية النسبية
4	22.5 – 20.5	3	0.12 = 25/3	9	0.36 = 25/9
5	24.5 – 22.5	7	0.28 =25/7	16	0.64 =25/16
6	26.5 – 24.5	4	0.16 = 25/4	20	0.80 =25/20

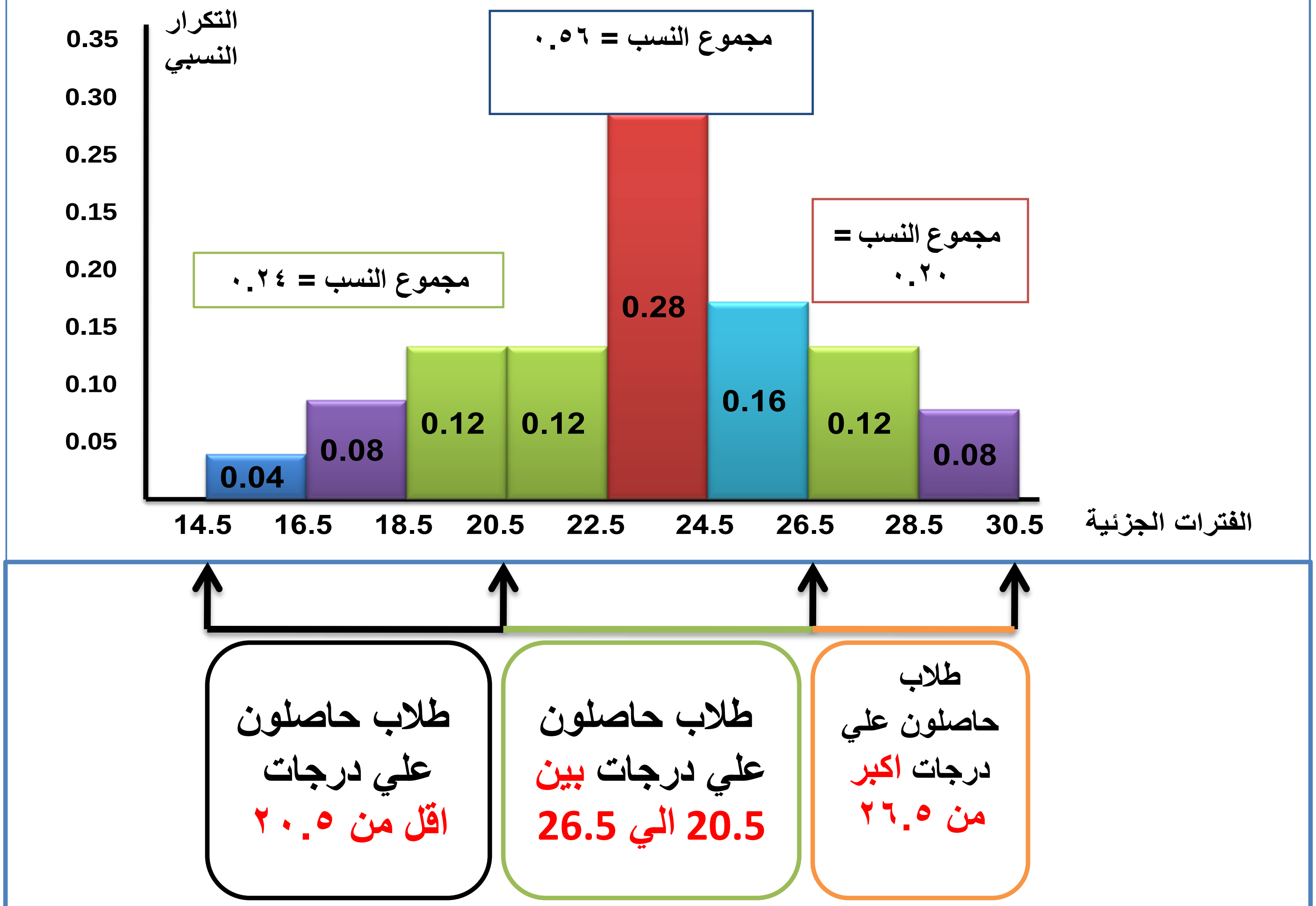
بنفس الطريقة السابقة نجد ان العمود المظلل باللون الاحمر يمثل مجموعة الطلاب الذين حصلوا على درجات بين 20.5 و 26.5

وبالتالي فإن نسبة الطلبة هي مجموع التكرارات النسبية الممثلة بالعمود المظلل باللون الاخضر

$$0.12 + 0.28 + 0.16 = 0.56$$

اما اذا استخدمنا المدرج التكراري النسبي بدلآ من الجدول التكراري النسبي

شكل (2) المدرج التكراري النسبي



اذا استبدلنا لفظ **نسبة** بلفظ **عدد** في الامثلة السابقة يفضل استخدام المدرج التكراري او الجدول التكراري بدلاً من المدرج التكراري النسبي و الجدول التكراري النسبي .

وبنفس الطريقة السابقة يمكن الاجابة علي الأسئلة الأتية:

احسب :

١- احتمال ان القيمة المختارة عشوائيا من البيانات تكون اقل من 20.5 .

٢ - احتمال وقوع القيمة المختارة عشوائيا من البيانات داخل الفترة 20.5إلى 26.5.

٣ - احتمال أن القيمة المختارة عشوائيا من البيانات تكون اكبر من 26.5.

الاجابة: (نفس الاجابة في الحالات الثلاثة السابقة)

لاحظ ان لفظ **نسبة** او **احتمال** لهما نفس المعنى.

تمت
بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ – ١٤٣١هـ

المحاضرة السادسة

عناصر المحاضرة

١-١- الطرق العددية لوصف البيانات

١-٢- القياسات العددية للنزعة المركزية

(اولآ- الوسط الحسابي)

1-1 الطرق العددية لوصف البيانات

تذكر

١- من أهم مميزات الطريقة البيانية لوصف القياسات هي عملية التمثيل المرئي للبيانات .

٢- في كثير من الأوقات نحتاج إلى تقرير كتابي عن البيانات ، وفى هذه الحالة <u>لا يمكن</u> استخدام الطريقة البيانية للحصول على هذا النوع من التقارير .

٣- وبالتالي فإننا نحتاج إلى الطريقة العددية لوصف القياسات وهى تتمثل في الحصول على

مجموعة أو فئة من الأعداد

تصف أو تصور التوزيع التكراري لهذه القياسات وتستخدم في نفس الوقت لصناعة الاستدلال أو الاستقراء حول خاصية معينة بالمجتمع .

٢-١ القياسات العددية للنزعة المركزية

مقياس النزعة المركزية (مركز التوزيع التكراري) هو قيمة تتوسط البيانات.

وتوجد طرق مختلفة لتعيين مركز مجموعة من القياسات وعلية توجد تعريفات مختلفة لمقياس النزعة المركزية اهمها، على سبيل المثال:

الوسط، الوسيط، المنوال، ونصف المدى

اولآ... الوسط الحسابي

- الوسط الحسابي لمجموعة تحتوى على n من البيانات

$y_1, y_2, y_3, ... , y_n$

يعرف على أنة مجموع هذه البيانات مقسوما على عددها (n).

- الرمز $\bar{y}$ يستخدم للدلالة على الوسط الحسابي للعينة بينما μ (يقرأ ميوه) يستخدم ليشير إلى الوسط الحسابي للمجتمع .

- من أهم استخدامات الوسط الحسابي للعينة $\bar{y}$ في مجال صناعة الاستدلال الإحصائي هو استخدامه كتقدير للوسط الحسابي للمجتمع μ.

-باستخدام رمز التجميع ، يمكن تعريف متوسط العينة بالصورة التالية :

$$\bar{y} = \frac{\sum_{i=1}^{n}(y_i)}{n}$$

تذكر

$$\bar{y} = \frac{\sum_{i=1}^{n}(y_i)}{n} = \frac{y_1 + y_2 + y_3 + \dots\dots\dots + y_n}{n}$$

مثال ١

اوجد الوسط الحسابي لمجموعة البيانات 2, 5, 7, 10, 11, 13

الحل:

بما أن عدد القياسات n هو 6 ، نحسب

$$\bar{y} = \frac{\sum_{i=1}^{n=6}(y_i)}{n=6} = \frac{y_1 + y_2 + y_3 + y_4 + y_5 + y_6}{6}$$

أي أن:

$$\bar{y} = \frac{\sum_{i=1}^{n=6}(y_i)}{n=6} = \frac{2+5+7+10+11+13}{6} = \frac{48}{6} = 8$$

لاحظ

القياسات في مثال ١ متساوية في عدد مرات ظهورها (تكرارها)

(كل قياس ظهر مرة واحدة فقط)

مثال ٢

أوجد الوسط لمجموعة البيانات

0.81, 0.8, 0.81, 0.81, 0.82, 0.81, 0.82, 0.8, 0.82, 0.81

<u>الحل</u> :

بما أن عدد القياسات n هو ١٠ ، نحسب

$$\bar{y} = \frac{\sum_{i=1}^{n=10}(y_i)}{n=10} = \frac{y_1 + y_2 + y_3 + y_4 + y_5 + \ldots\ldots\ldots + y_9 + y_{10}}{10}$$

أي أن :

$$\bar{y} = \frac{\sum_{i=1}^{n=10}(y_i)}{n=10} = \frac{0.81+0.8+0.81+0.81+0.82+0.81+0.82+0.8+0.82+0.81}{10} = \frac{8.11}{10} = 0.811$$

لاحظ

يمكن أن تصبح العملية الأخيرة (عملية الجمع) أسرع وأكثر سهولة إذا تم تجميع الأرقام المتشابهة معا كما يلي :

$$\bar{y} = \frac{\sum_{i=1}^{n=10}(y_i)}{n=10} = \frac{(0.8+0.8)+(0.81+0.81+0.81+0.81+0.81)+(0.82+0.82+0.82)}{10}$$

$$= \frac{(2)\,x\,(0.8)+(5)\,x\,(0.81)+(3)\,x\,(0.82)}{10}$$

حيث ٢ هي عدد مرات ظهور (تكرار) القيمة 0.8

حيث 5 هي عدد مرات ظهور (تكرار) القيمة 0.81

حيث 3 هي عدد مرات ظهور (تكرار) القيمة 0.82

لاحظ

القياسات في مثال ٢ غير متساوية في عدد مرات ظهورها (تكرارها)

(كل قياس ظهر عدد مرات مختلف)

وبالتالي اذا كانت:

القيمة	y_1	y_2	y_3	....	....	y_n
تكرارها	f_1	f_2	f_3	....	....	f_n

فإن الوسط الحسابي يعطي بالشكل:

$$\bar{y} = \frac{\sum_{i=1}^{n} y_{i.} f_{i}}{n} = \frac{y_{1.} f_{1} + y_{2.} f_{2} + \dots + y_{n.} f_{n}}{n}$$

مجموع حاصل ضرب كل قيمة في تكرارها

مقسوماً علي عدد القيم (مجموع التكرارات).

ملاحظة

-في المثالين السابقين اعطيت جميع القياسات (البيانات) بطريقة (صريحة) (تم سرد جميع القياسات) إما بدون تكرار (مثال ١) او بتكرارات مختلفة (مثال ٢).

- لإيجاد الوسط الحسابي لمجموعة من البيانات المبوبة (معطاة بطريقة ضمنية) خلال جدول تكراري بطول فترة جزئية محددة ، فإن أسهل الطرق لإتمام هذه العملية هو تصميم جدول يلخص خطوات أو طريقة الحل (كما في المثال التالي):

مثال ٣: أوجد الوسط الحسابي لمجموعة البيانات المبوبة في الجدول التالي:

حدود الفترات	التكـــــرار
171-175	٤
166-170	٨
161-165	١٤
156-160	٢٢
151-155	٢٧
146-150	١٩
141-145	١٧
136-140	١١
131-135	٣

الحل: بفرض أن x تمثل مركز الفترة الجزئية (أو الفصل)، f تمثل تكرار هذه الفترة الجزئية (أو الفصل).

الفترة i	حدود الفترات	مركز الفترة x_i	التكـــــرار f_i	$x_i f_i$
1	171-175	(١٧١+١٧٥) / ٢ = ١٧٣	٤	692
2	166-170	(١٦٦+١٧٠) / ٢ = ١٦٨	٨	1344
3	161-165	(١٦١+١٦٥) / ٢ = ١٦٣	١٤	2282
4	156-160	(١٥٦+١٦٠) / ٢ = ١٥٨	٢٢	3476
5	151-155	(١٥١+١٥٥) / ٢ = ١٥٣	٢٧	4131
6	146-150	(١٤٦+١٥٠) / ٢ = ١٤٨	١٩	2812
7	141-145	(١٤١+١٤٥) / ٢ = ١٤٣	١٧	2431
8	136-140	(١٣٦+١٤٠) / ٢ = ١٣٨	١١	1518
9	131-135	(١٣١+١٣٥) / ٢ = ١٣٣	٣	399
المجموع	....	....	١٢٥	19085

من تعريف الوسط الحسابي نجد ان:

$$\bar{x} = \frac{\sum_{i=1}^{9} x_i \cdot f_i}{n} = \frac{19085}{125} = 152.68$$

تمت

بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ – ١٤٣١هـ

المحاضرة السابعة

عناصر المحاضرة

١-١ تابع القياسات العددية للنزعة المركزية

(الوسيط – المنوال – نصف المدى)

٢-١ مميزات وعيوب مقاييس النزعة المركزية

1-1 تابع القياسات العددية للنزعة المركزية

ثانياً ... الوسيط

١- **الوسيط** أو **القيمة الوسيطية** لمجموعة من البيانات عددها n

$$y_1, y_2, y_3, \dots, y_n$$

هي تلك القيمة y التي **تتوسط** القيم بعد **ترتيبها تبعا لقيمها** .

٢- إذا كان عدد البيانات في المجموعة **زوجيا** فإنه توجد **قيمتان** تتوسطان هذه البيانات ويكون الوسيط في هذه الحالة هو **الوسط الحسابي لهاتين القيمتين**. (كما بالشكل التالي)

عدد البيانات فردي

١ ٢ ٣ ٤ ٥

قيمتان علي اليمن

الوسيط

قيمتان علي اليسار

بينما

عدد البيانات زوجي

١ ٢ ٣ ٤ ٥ ٦

قيمتان علي اليمن

الوسيط = (٣+٤)/٢

قيمتان علي اليسار

مثال ١

اوجد **الوسيط** لكل مجموعة من القياسات التالية :

١- 5 , 3 , 2 , 7 , 4

٢- 10 ,8 ,13 , 14 , 9 , 8

الحل:

١- **أولا** : **بترتيب** القياسات **تبعا لقيمها** نجد أنها تأخذ الصورة :

2 , 3 , 4 , 5 , 7

ثانيا : نجد أن **الوسيط أو القيمة المتوسطة هي 4** .

٢- **أولا** : **بترتيب** القياسات **تبعا لقيمها** نجد أنها تأخذ الصورة :

14 , 13 , 10 , 9 , 8, , 8

ثانيا : وحيث أن عدد قياسات هذه المجموعة n هي 6 (**زوجي**)

فإن **الوسيط أو القيمة المتوسطة** تعطى بالشكل :

$$\frac{9+10}{2} = 9.5$$

مثال ٢

أوجد **الوسط الحسابي** و**الوسيط** لمجموعة من القياسات التالية :

14 , 13 , 11 , 10 , 10, 8 , 7 , 5

الحل:

اولآ- لإيجاد **الوسط الحسابي** $\bar{y}$ ، نحسب $\sum_{1=1}^{8} y_i = 78$

ثم نوجد

$$\bar{y} = \frac{\sum_{i=1}^{8} y_i}{n} = \frac{78}{8} = 9.75$$

ثانياً: لإيجاد الوسيط أو القيمة الوسيطية للقيم

14 , 13 , 11 , 10 , 10, 8 , 7 , 5

نلاحظ أن القيم مرتبة ولا تحتاج إلى إعادة ترتيب وكذلك نلاحظ أن عددها زوجي وبالتالي نجد أن الوسيط لهذه المجموعة هو

$$\frac{10+10}{2}=10$$

ثالثاً... المنوال

١- المنوال لمجموعة من البيانات هي القيمة صاحبة أكبر تكرار.

٢- إذا احتوت المجموعة على قيمتين لهما أكبر تكرار فإنه يقال أن مجموعة البيانات في هذه الحالة ثنائية المنوال.

٣- أما إذا احتوت مجموعة البيانات على أكثر من قيمتين لهم أكبر تكرار فإنه يقال أن مجموعة البيانات في هذه الحالة متعددة المنوال.

٤- إذا لم تحتو مجموعة البيانات على أي تكرار فإن المجموعة في هذه الحالة لا تحتوى على منوال (لا يوجد منوال).

مثال ١

أوجد **المنوال** لكل مجموعة من القياسات التالية :

1- 5 , 3 , 4 , 1 , 5 , 1 , 3 , 5 , 5 , 5

2 - 9 , 7 , 6 , 6 , 6 , 5 , 4 , , 3 , 2 , 2 , , 2 , 1

٣- 1 , 2 , 3 , , 6 , 7 , 8 , 9 , 10

الحل :

١- العدد 5 هو المنوال حيث أن له أكبر تكرار (5 مرات) .

٢- العددان 6 , 2 يمثلان المنوال حيث أن لكل منهما أكبر تكرار (3 مرات) .

٣- لا يوجد منوال لهذه المجموعة حيث أنة لا يوجد تكرار لأي من القياسات .

رابعاً... نصف المدى

نصف المدى هو القيمة التي تتوسط أكبر وأصغر القياسات .

$$\text{نصف المدى} = \frac{\text{اكبر قيمة للقياسات} + \text{اصغر قيمة للقياسات}}{٢}$$

1-2 مميزات وعيوب مقاييس النزعة المركزية

١- تعتبر عملية اختيار أكثر الطرق ملائمة لقياس النزعة المركزية عملية غير بسيطة ولا توجد طريقة معينة لتحديد أكثر الطرق ملائمة لجميع أنواع البيانات.

٢- طرق القياس المختلفة للنزعة المركزية لها الكثير من المميزات والعيوب والجدول التالي يوضح أهم مميزات وعيوب القياسات المختلفة للنزعة المركزية .

المعدل (المتوسط)	تعريفه	مدى استخدامه	إيجاده	تأثره بالقيم المتطرقة	تأثره بجميع القيم	مميزاته وعيوبه
الوسط الحسابي	$\bar{x} = \dfrac{\sum x}{n}$	الأكثر استخداما	دائما	نعم	نعم	يعمل بكفاءة مع جميع الطرق الإحصائية
الوسيط	القيمة التي تتوسط القيم	غالبا	دائما	لا	لا	غالبا ما يستخدم في حالة وجود قيم متطرقة
المنوال	القيم الأكثر تكرارا	أحيانا	أحيانا لا يوجد وأحيانا أكثر من واحد	لا	لا	صالح للبيانات من المستوى الاسمي
نصف المدى	$\dfrac{\text{الأكبر} + \text{الأصغر}}{2}$	نادرا	دائما	نعم	لا	يتأثر بالقيم المتطرقة

تمت
بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ ـ ١٤٣١هـ

المحاضرة الثامنة

عناصر المحاضرة

١-١- مقاييس الاختلاف أو التشتت

- المدي
- المئين
- الربيعات

مقدمة

١- في البند السابق تعرفنا على طرق إيجاد

مقاييس النزعة المركزية

والآن نتعرف على مقاييس

الاختلاف أو التغير أو التشتت

لمجموعة البيانات .

٢- تأتى أهمية مقاييس الاختلاف لعدم كفاءة مقاييس النزعة المركزية وحدها لوصف مجموعة البيانات (كما بالمثال التالي).

مثال: بفرض المجموعتين الآتيتين من البيانات:

مقاييس النزعة المركزية	المجموعة الثانية	مقاييس النزعة المركزية	المجموعة الاولى
الوسط = ١٠ الوسيط = ١٠ نصف المدي = ١٠ المنوال : لا يوجد	$x_1 = 1$ $x_2 = 10$ $x_3 = 19$	الوسط = ١٠ الوسيط = ١٠ نصف المدي = ١٠ المنوال : لا يوجد	$y_1 = 9$ $y_2 = 10$ $y_3 = 11$

كل من المجموعتين لهما نفس **مقاييس النزعة المركزية**، على الرغم من أن المجموعة **الثانية** تحتوى على **اختلاف** (بعد قيم *المجموعة عن الوسط الحسابي*) **أكبر** من المجموعة **الأولى** .

قد يكون **الاختلاف** أو **التغير** **مطلوبا أحيانا** في بعض البيانات فعلى سبيل المثال فإن:

الشركات المنتجة لقطع الغيار **لاتحتاج** إلى **تغير** أو **اختلاف** في المنتج الواحد والذي يشير إلى **عدم جودة الإنتاج** ،

بينما يتغير الحال في أي

اختبار تربوي لاختيار بعض العاملين في مجال التدريس ، حيث أننا **نحتاج** إلي **الاختلاف** أو **التغير** في نتائج الاختبار والذي **يتيح المجال الأكبر** لعملية **الاختيار الصحيحة**.

المدى

سبق أن تعرفنا على أسهل الطرق لقياس **الاختلاف** ، وهو **المدى** ، والذي يعرف على أنة **الفرق بين أكبر وأصغر البيانات**.

اوجد المدى لكل من:

أ)المجموعة

74	34	73	23
17	26	29	28
49	52	8	88
45	32	96	37
62	23	62	81

المدى = 96 - 8 = 88

ب)المجموعة

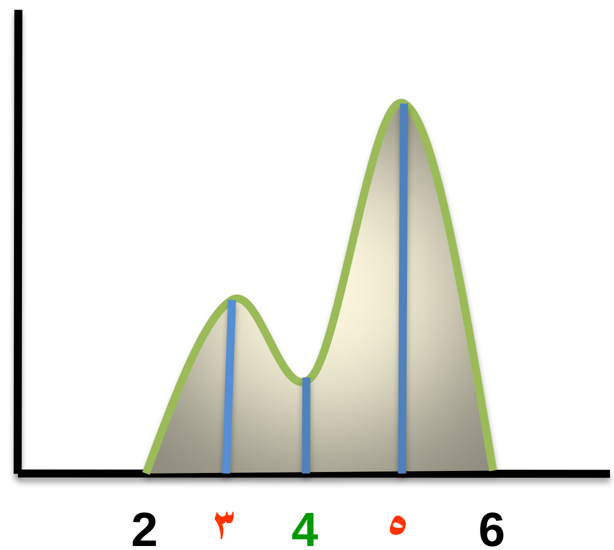
8.8	6.7	7.1	2.9
9.0	0.2	1.2	8.6
6.3	4.6	2.1	8.8

المدى = 9 – 0.2 =8.8

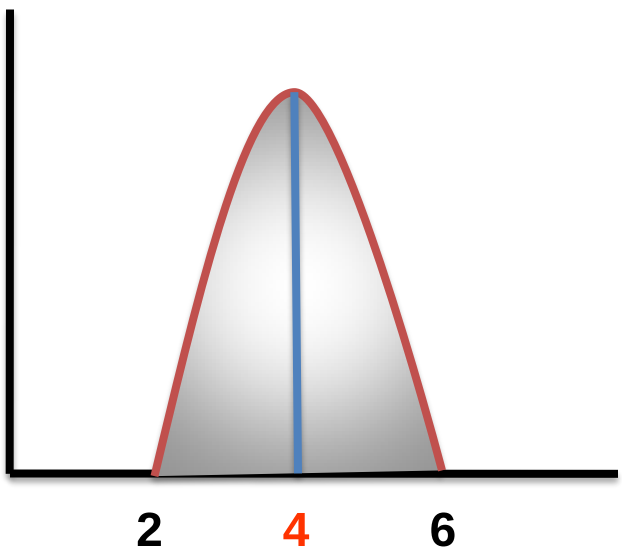
لاحظ

الشكلان الآتيان (٢،١) يوضحان توزيعين لمجموعتين من البيانات وبالمقارنة نجد أن المجموعتين رغم أن لهما نفس المدى إلا أن الاختلاف بين قياسات كل مجموعة لايساوى الاختلاف بين قياسات المجموعة الأخرى ، مما يؤكد أن

المدى لا يكفى وحدة كمقياس لعملية الاختلاف أو التغير.



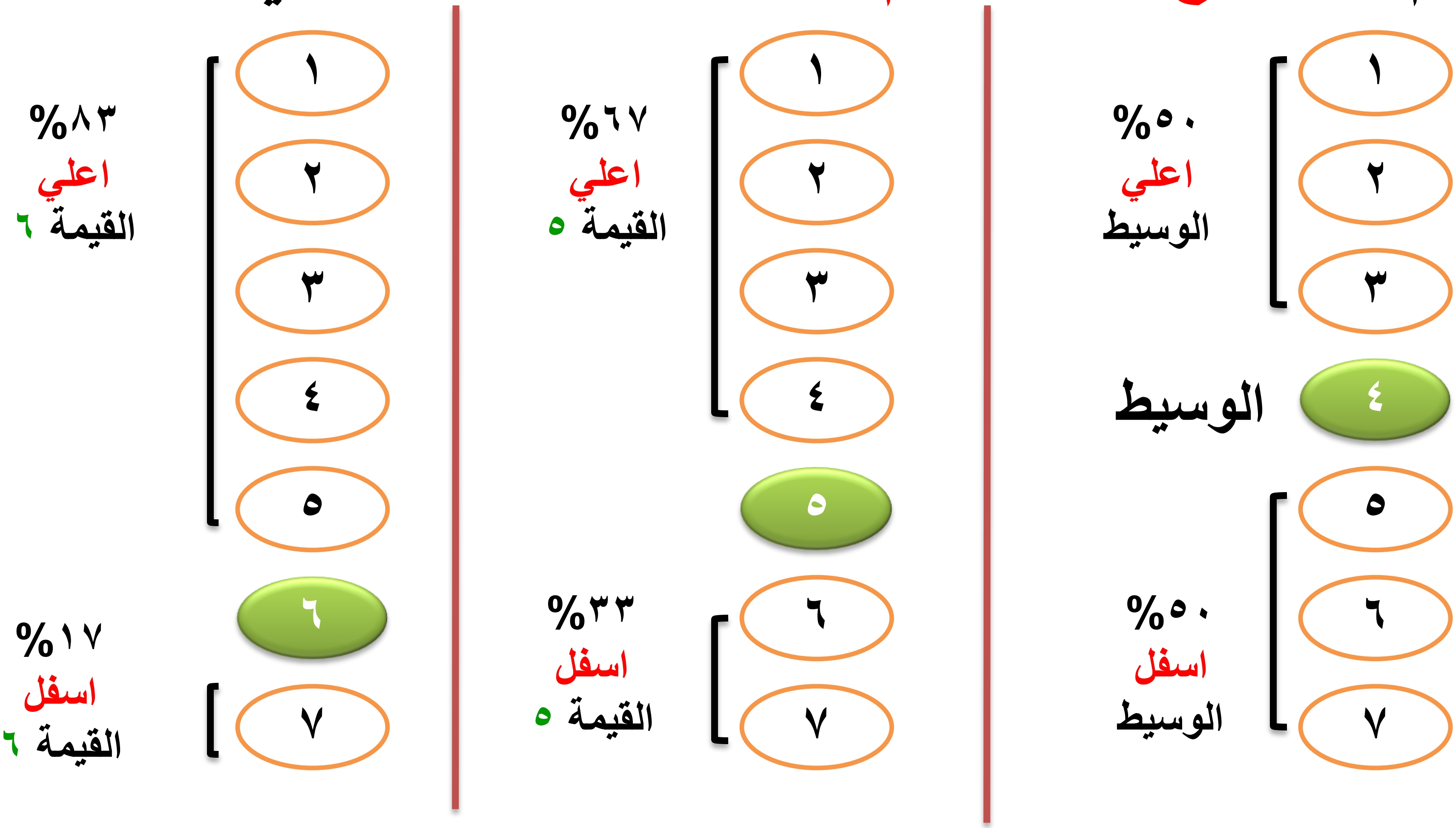
شكل ١



شكل ٢

المئين

للحصول على مقياس أكثر دقة من المدى لقياس الاختلاف فإننا نقوم بتوسيع أو بتعميم فكرة الوسيط كما يلي:



وبالتالي

يمكن تعريف المئين كما بالشكل التالي

y_1

y_2

y_3

y_4

y_5

....

y

....

....

y_n

(100 – P)%
اعلي (فوق)
y

المئين p

P %
اسفل (تحت)
y

أي أن

اذا كان لدينا مجموعة تحتوى على n من البيانات

$$y_1, y_2, y_3, \ldots\ldots, y_n$$

منظمة أو مرتبة تبعا لقيمها.

فإنه يمكن تعريف المئين P على إنه تلك القيمة y بحيث أن **p%** من القياسات (تقع تحتها) و **% (100-p)** (تقع فوقها) .

-يعتبر **المئين**

أكثر دقة وأكثر حساسية

من **المدى** كمقياس للتغير أو الاختلاف ولكن **أكثر عيوبه** أننا **نحتاج إلى**

العديد من المئينات

لوصف مجموعة البيانات وصفاً كافيا.

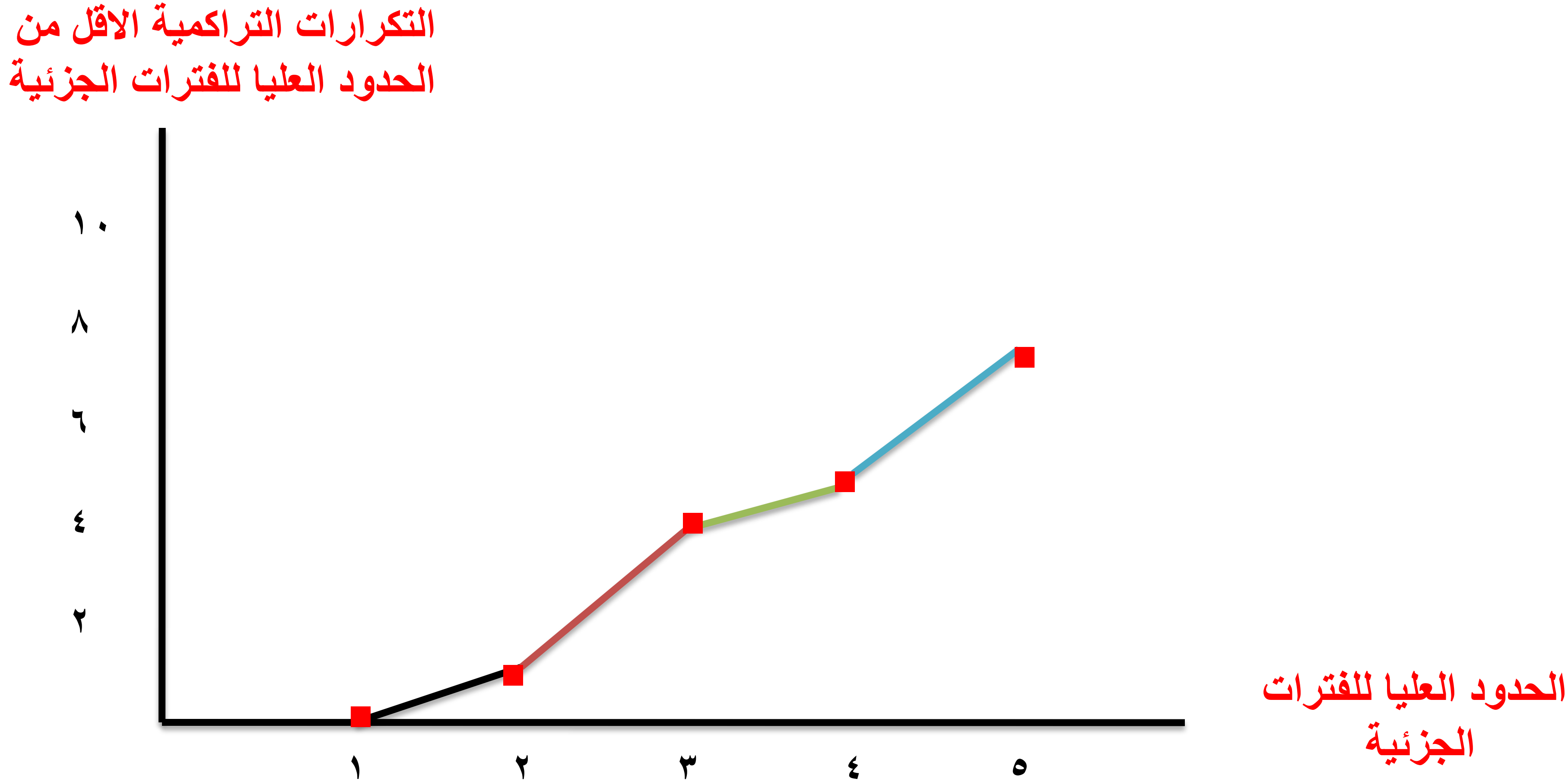
-**المنحنى الخطى** ويسمى

المضلع التكراري التراكمي

ويمكن الحصول علية كما يلي:

-رسم **التكرار التراكمي** الأقل **من الحدود العليا للفترات الجزئية** ضد **الحدود العليا للفترات الجزئية**،

- ثم توصيل **النقاط المتتابعة بخط مستقيم** (انظر الشكل) .



المضلع التكراري التراكمي

عند استخدام التكرارات النسبية بدلا من التكرارات العادية (المطلقة) فإن المنحنى يسمى

المضلع التكراري التراكمي النسبي
أو
المضلع التكراري التراكمي المئوي

والذي يمتاز بسهولة وسرعة الحصول على المعلومة من خلاله.

مثال: الجدول التالي يوضح درجات عينة مكونة من **40** طالب :

2.2	4.1	3.5	4.5	3.2	3.7	3.0	2.6
3.4	1.6	3.1	3.3	3.8	3.1	4.7	3.7
2.5	4.3	3.4	3.6	2.9	3.3	3.9	3.1
3.3	3.1	3.7	4.4	3.2	4.1	1.9	3.4
4.7	3.8	3.2	2.6	3.9	3.0	4.2	3.5

أ- أوجد p_{80}(المئين 80) لتوزيع درجات الطلاب .

ب- أوجد p_{85}(المئين 85) لتوزيع درجات الطلاب .

ج- ارسم المضلع التكراري التراكمي النسبي ثم اوجد باستخدامه كلا من المئين 25 والمئين 70.

الحل: اولآ- نكون الجدول التكراري (جدول ١)

التكرار التراكمي النسبي	التكرار التراكمي	التكرارات النسبية	التكرارات المطلقة f_i	حدود الفترة	الفترة الفصل *i*
2/40	2	2/40 = .05	2	1.45 – 1.95	1
3/40	3	1/4 = .30	1	1.95 – 2.45	2
7/40	7	4/40 = .1	4	2.45 – 2.95	3
22/40	22	15/40 = .38	15	2.95 – 3.45	4
32/40	32	10/40 = .25	10	3.45 – 3.95	5
37/40	37	5/40 = .13	5	3.95 – 4.45	6
40/40=1	40	3/40 = .08	3	4.45 – 4.95	7

نستخدم الاعمدة الثلاثة الملونة كما يلي (علي حسب المطلوب)

اما العمود باللون الاخضر (حدود الفترات) مع العمود باللون الاحمر (المضلع التكراري)

او العمود باللون الاخضر (حدود الفترات) مع العمود باللون الاصفر (المضلع التكراري النسبي)

ثانياً-نكون الجدول التالي

(نستخدم الاعمدة الثلاثة الملونة من (جدول ١) ونستخدم الحدود العليا الموضحة باللون الاحمر في عمود (حدود الفترات) نحصل علي :

الفترات الجزئية	التكرار التراكمي	نسبة التكرار التراكمي
أقل من 1.45	0	(0/40) .100 = 0
اقل من 1.95	2	(2/40) .100 = 5.0
اقل من 2.45	3	(3/40) .100 = 7.5
اقل من 2.95	7	(7/40) .100 = 17.5
اقل من 3.45	22	(22/40) .100 = 55
اقل من 3.95	٣٢	(32/40) .100 = ٨٠
اقل من 4.45	37	(37/40) .100 = 92.5
اقل من 4.95	40	(40/40) .100 = 100

والان

أ- لحساب قيمة p_{80} (المئين 80) فإننا نبحث عن القيمة

(الدرجة) y في العينة والتي تقع تحتها $\frac{80}{100} x 40 = 32$ من القيم

أوالدرجات في العينة.

من (جدول ٢) يوجد ٨٠% من البيانات (في عمود نسبة التكرار التراكمي) اي

32 قيمة (في عمود التكرارات) تقع تحت الحد 3.95(في عمود الفترات الجزئية).

وبالتالي فإن p_{80} (المئين 80) = 3.95.

لكن

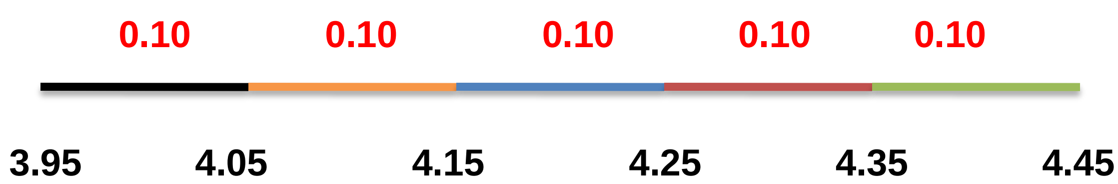
ب- لحساب قيمة p_{85} (المئين 85) فإننا نبحث عن القيمة (الدرجة) y في

العينة والتي تقع تحتها $\frac{85}{100} x 40 = 34$ من القيم أوالدرجات في العينة.

يوجد 32 قيمة تقع تحت الحد 3.95 (كماهوموضح بالفقرة أ). ولاستكمال هذا العدد من القيم فإننا

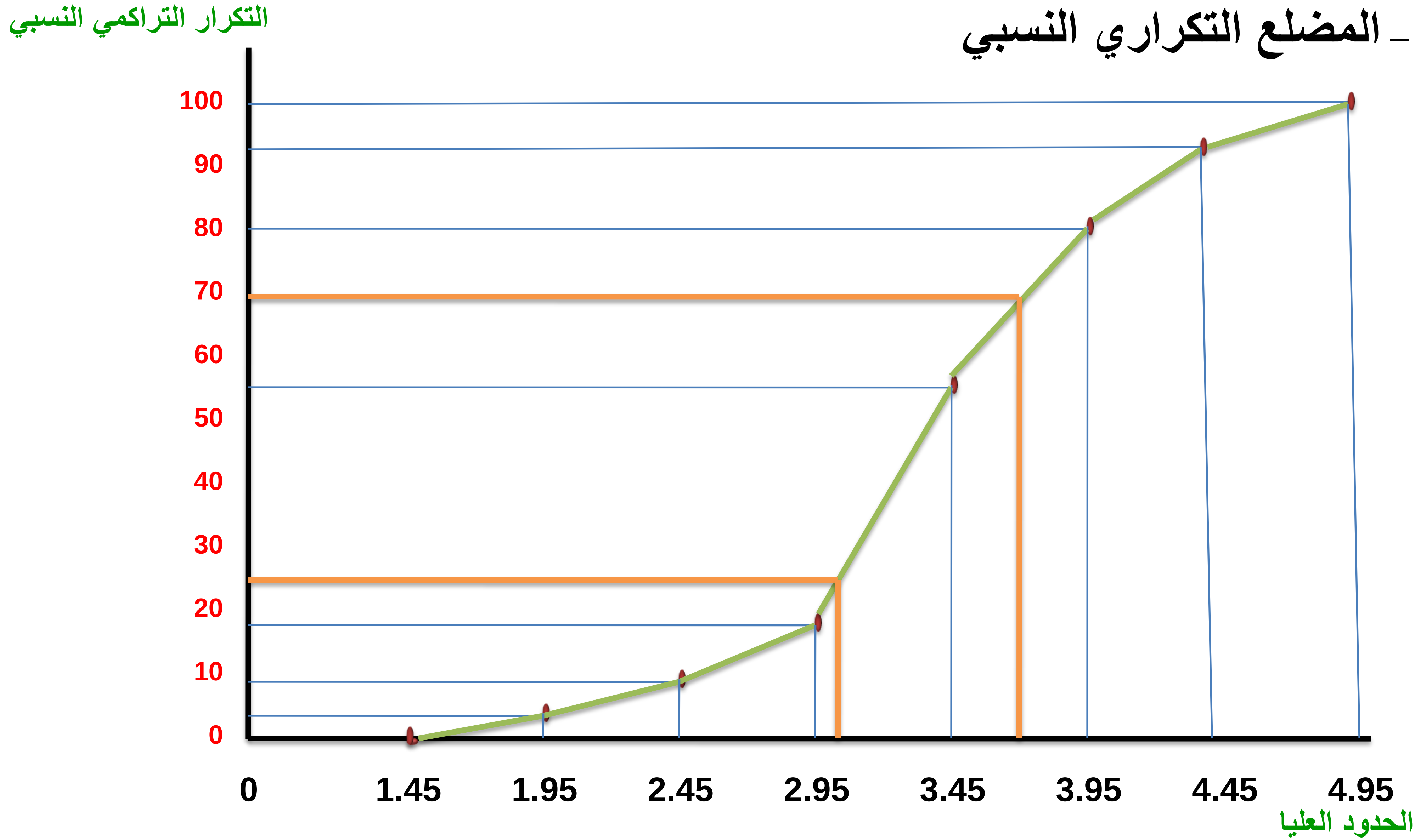
نحتاج إلى 2 من 5 قيم موزعة بين الحدين 3.95 ، 4.45 وبالتالي نتحرك مسافة $\frac{2}{5} x 0.5 = 0.2$

بعد الحد 3.95 وعلية تكون قيمة



$$p_{85} = 3.95 + 0.2 = 4.15$$

ج ـ المضلع التكراري النسبي



الربيعات

١- المئين p_{25} (المئين ٢٥) يسمى الربيع الأدنى (أوالأول) ويرمز له بالرمز Q_1

أي أن $p_{25} = Q_1$

٢- المئين p_{75} (المئين ٧٥) يسمى الربيع الأعلى (أوالثالث) ويرمز له بالرمزQ_3

أي أن $p_{75} = Q_3$

٣- المئين p_{50} (المئين ٥٠) يسمي الربيع الأوسط (أو الثاني) أو الوسيط

ويرمز له بالرمز Q_2 أي أن $p_{50} = Q_2$.

تمت
بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧

مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ ـ ١٤٣١هـ

المحاضرة التاسعة

عناصر المحاضرة

١-١- مقاييس الاختلاف أو التشتت (تابع المحاضرة الثامنة)

- الإنحراف
- التباين
- الإنحراف المعياري

الإنحراف

- سنقوم الآن بعرض مقياس هام **للتغير** أو **الاختلاف** معتمدين على **تشتت أو انحراف القياسات عن الوسط الحسابي للعينة.**

انحراف القيمة y_1 (بعد القيمة y_1 عن الوسط $\bar{y}$)

$$y_1 \text{————————————} \bar{y}$$

$$(y_1 - \bar{y})$$

- بفرض أن $(y_i - \bar{y})$ تمثل **انحراف القيمة رقم** i عن الوسط الحسابي.

- **الانحرافات الكبيرة** تشير عادة إلى **تغيرات** أو **اختلافات كبيرة**.

مثال

انحراف القيمة (٩) عن الوسط (١٠)
هي
9-10 = -1

٩

١٠ الوسط

١١

انحراف القيمة (١١) عن الوسط (١٠)
هي
11-10 = +1

نلاحظ ان **مجموع الإنحرافات عن الوسط** = (10)

انحراف القيمة (٩) عن الوسط (١٠) + انحراف القيمة (11) عن الوسط (١٠)

أي أن = (-1) + (+1) = صفر.

وبالتالي فإن <u>مجموع n من الانحرافات عن الوسط الحسابى يساوى صفر</u> . أي أن

$$\sum_{i=1}^{n}(y_i - \bar{y}) = 0$$

لمنع حدوث المجموع الصفري نقوم **بتربيع** الإنحرافات عن الوسط

أي نوجد $\left(y_i-\bar{y}\right)^2$ (عن طريق **ضرب** مقدار الإنحراف في **نفسه**).

في المثال السابق : الإنحرافات عن الوسط كانت (-1) و (+1)

وبالتالي مربعات الإنحرافات عن الوسط تصبح (+1) و (+1)

وعليه يكون **مجموع** مربعات الإنحرافات عن الوسط = (+1) + (+1) = 2+ (لا تساوي الصفر)

وبالتالي يكون $\sum_{i=1}^{n}\left(y_i-\bar{y}\right)^2$ لا يساوي الصفر.

- مجموع مربعات الانحرافات يعتبر مقياسا **للإنتشار**.
- **يعرف (متوسط مربع الانحرافات في المجتمع) بتباين المجتمع** ويرمز له بالرمز σ^2 (سيجما تربيع).

التباين

تعريف

بفرض أن لدينا مجموعة تحتوى على n من القياسات

$y_1, y_2, \ldots\ldots\ldots, y_n$

فإن تباين العينة يرمز له بالرمز S^2 ويعطى بالمعادلة:

$$s^2 = \frac{1}{n-1} \sum_{i-1}^{n} \left(y_i - \bar{y}\right)^2$$

مثال ١ احسب الوسط الحسابي والتباين للعينة 6 ,5 ,3 ,2 ,4.

الحل: أولاً- نحسب الوسط الحسابي كما يلي:

$$\bar{y}=\frac{\sum_{i=1}^{5} y_i}{5}=\frac{20}{5}=4$$

ثانياً- لحساب تباين العينة نوجد الجدول التالي:

y_i	$y_i-\bar{y}$	$(y_i-\bar{y})^2$
4 2 3 5 6	(4 – 4)= 0 (2 – 4)= -2 (3 – 4)= -1 (5 – 4)= 1 (6 – 4)= 2	0 4 1 1 4
$\sum y_i=20$	$\sum(y_i-\bar{y})=0$	$\sum(y_i-\bar{y})^2=10$

وبالتالي

$$s^2=\frac{1}{n-1}\sum_{i-1}^{n}(y_i-\bar{y})^2=\frac{1}{5-1}\sum_{i-1}^{5}(y_i-\bar{y})^2=\frac{10}{4}=2.5$$

طريقة مختصرة لحساب التباين

يمكننا إستخدام المعادلة التالية لحساب قيمة التباين

$$s^2 = \frac{1}{n-1}\left(\sum_{i=1}^{n} y_i^2 - \frac{\left(\sum_{i=1}^{n} y_i\right)^2}{n}\right) \quad \text{.......} \quad \text{.......} \ (*)$$

-نلاحظ في هذه المعادلة اننا لابحتاج الي حساب قيمة الوسط الحسابي .

- هذه العملية الحسابية تحتاج إلى :

١- المجموع الحسابي العادي للقياسات $\sum_{i=1}^{n} y_i$

٢- مجموع مربعات القياسات $\sum_{i=1}^{n} y_i^2$

مثال ٢

احسب S^2 للعينة في مثال (١)

الحل : نكون الجدول التالي:

y_i	y_i^2
4	16
2	4
3	9
5	25
6	36
$\sum y_i = 20$	$\sum y_i^2 = 90$

بالتعويض في المعادلة (*) نحصل علي :

$$s^2 = \frac{1}{5-1}\left(90 - \frac{(20)^2}{5}\right) = \frac{1}{5-1}(90-80) = \frac{10}{4} = 2.5$$

الإنحراف المعياري

- **الانحراف المعياري** للعينة التي حجمها n ، يرمز له بالرمز S ، يعرف على إنه الجذر التربيعي الموجب لتباين العينة s^2 أي أن

$$s = \sqrt{s^2} = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \bar{y})^2}{n-1}}$$

في المثال السابق الإنحراف المعياري

$$s = \sqrt{s^2} = \sqrt{2.5} = 1.58$$

تمت
بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ ـ ١٤٣١هـ

المحاضرة العاشرة

الباب الثاني – مقدمة في الاحتمالات

عناصر المحاضرة

- التجربة العشوائية.
- فراغ العينة.
- الاحداث وانواعها.
- تعريف الاحتمال .

مقدمة

تعطى نظرية الاحتمالات الأساس الرياضي لعملية الاستدلال أو الاستنتاج من المجتمع عن العينة المأخوذة من المجتمع .

أما علم الإحصاء فهو دراسة كيفية استخدام المعلومات الموجودة في العينة للاستدلال أو الاستنتاج عن المجتمع .

١- التجربة هي عملية الحصول على معلومة أو بيان .

٢- نتائج التجربة تسمى أحداث .

٣- الأحداث إما أن تكون بسيطة أو مركبة .

5- الحدث يكون بسيطا إذا لم يمكن تحليلية .

6- الأحداث غير البسيطة تكون مركبة .

نقاط العينة

١- يجب وضع طريقة لتصور وسرد **الأحداث البسيطة** أو **نقاط العينة** للتجربة المعطاة .

2- فئة **جميع نقاط العينة** للتجربة يرمز لها لها بالرمز **S** وتسمى **فراغ العينة**.

3- يمكن إعادة تعريف **الحدث** على أنة **فئة جزئية من فراغ العينة**

4- تستخدم **أشكال فن** لتمثيل فراغ العينة **كنطاق محدود بمنحى** وتمثيل نقاط العينة **كنقاط داخل هذا النطاق** .

٥- بما ان كل مجموعة جزئية من نفسها فإن **S** (فراغ العينة) يكون حدثاً ويسمى **بالحدث المؤكد.**

٦- **المجموعة الخالية** (التي لا تحتوى على اي عناصر) هى مجموعة جزئية من **S** (فراغ العينة) وبالتالى تكون حدثاً ويسمى **بالحدث المستحيل.**

مثال ١

ألقي حجر نرد مرة واحدة – اكتب كلا من الأحداث التالية ثم وضح أى من هذه الأحداث يكون حدث مؤكد وأيهم يكون حدث مستحيل:

أ- حدث A الحصول على عدد زوجى.

ب – حدث B الحصول على عدد يقبل القسمة على ٣.

ج – حدث C الحصول على عدد اكبر من ٦.

د – حدث D الحصول على عدد اكبر من أو يساوى ١ وأقل من أو يساوى ٦.

الحل:

فراغ العينة هو $S = \{ 1, 2, 3, 4, 5, 6 \}$

أ- حدث **A** الحصول على عدد **زوجى** نعبر عنه بالمجموعة:

$$A = \{ 2, 4, 6 \}$$

ب - حدث **B** الحصول على عدد **يقبل القسمة على ٣** نعبر عنه بالمجموعة:

$$B = \{ 3, 6 \}$$

ج - حدث **C** الحصول على عدد **اكبر من ٦** نعبر عنه بالمجموعة:

المجموعة الخالية = $C = \{ \ \}$ (حدث مستحيل)

د - حدث **D** الحصول على **عدد اكبر من أو يساوى ١ وأقل من أو يساوى ٦.**

$D = \{ 1, 2, 3, 4, 5, 6 \} = S$ (حدث مؤكد)

مثال ٢

في تجربة لترتيب ثلاثة مرشحين M_1 وM_2 وM_3 لوظيفة ما تبعا لمقدرتهم على أداء هذه الوظيفة.

الحل: نقاط العينة لهذه التجربة يمكن أن يرمز لها باستخدام **الثلاثيات المرتبة** والتى يمكن الحصول عليها كما يلي:

المرشح الأول	المرشح الثاني	المرشح الثالث
M_1	M_2	M_3
	M_3	M_2
M_2	M_3	M_1
	M_1	M_3
M_3	M_1	M_2
	M_2	M_1

وبالتالي تكون نقاط العينة لهذه التجربة هي:

(M_3 M_2 M_1) ($M_{١}$ $M_{٣}$ $M_{٢}$) ($M_{٢}$ $M_{١}$ $M_{٣}$)

($M_{٢}$ $M_{٣}$ M_1) (M_3 $M_{١}$ $M_{٢}$) ($M_{١}$ M_2 $M_{٣}$)

والآن

الحدث المركب الذي يشير إلى أن المرشح الثاني M_2 يكون ترتيبه **الأول** يحتوى على **اثنين** من نقاط العينة هما

($M_{١}$ $M_{٣}$ $M_{٢}$)

(M_3 $M_{١}$ $M_{٢}$)

بينما

الحدث الذي يشير إلى أن المرشح الثاني M_2 يكون ترتيبه الأول والمرشح الأول M_1 ترتيبه الثاني هو **حدث بسيط** يحتوى على نقطة واحدة من نقاط العينة هي

(M_3 $M_١$ $M_٢$)

أهمية الإحتمالات

١ - بعد تحديد **فراغ العينة** تحديدا جيدا فإنه من الضروري **ربط الاحتمالات** بكل **نقاط العينة** حيث أن

احتمال أي حدث A يساوي مجموع احتمالات نقاط العينة داخله.

٢- الاحتمالات **المرتبطة بنقاط العينة** (الأحداث البسيطة) .... $E_1, E_2,$ يجب أن تحقق الشروط الآتية :

$$0 \le p(E_i) \le 1$$

و

$$\sum_i p(E_i) = 1$$

٣- إذا كان **عدد النتائج الممكنة لتجربة ما** يساوى N وكان **لكل** نتيجة **نفس الفرصة في الظهور** في فراغ العينة فإن المعادلة

$$\sum_{i=1}^{N} p(E_i) = 1$$

تتطلب أن يكون **الاحتمال المرتبط بكل نقطة من نقاط العينة مساويا** $\frac{1}{N}$.

وفى هذه الحالة إذا كان الحدث A يحتوى على n_A من نقاط العينة فإن <u>احتمال ظهور الحدث</u> A ويرمز لها $p\ (A)$ ويعرف كما يلي:

$$p(A) = \frac{n_A}{N}$$

وبالتالى

فإن **احتمال** **الحدث المركب** الذي يشير إلى أن المرشح الثاني M_2 يكون ترتيبه **الأول** والذى يحتوى على **اثنين** من نقاط العينة هما

$$(\; M_{١} \;\; M_{٣} \;\; \boxed{M_{٢}} \;)$$

$$(\; M_3 \;\; M_{١} \;\; \boxed{M_{٢}} \;)$$

يساوى $\frac{2}{6}$.

تمت

بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧

مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ ـ ١٤٣١هـ

المحاضرة الحادية عشر

عناصر المحاضرة

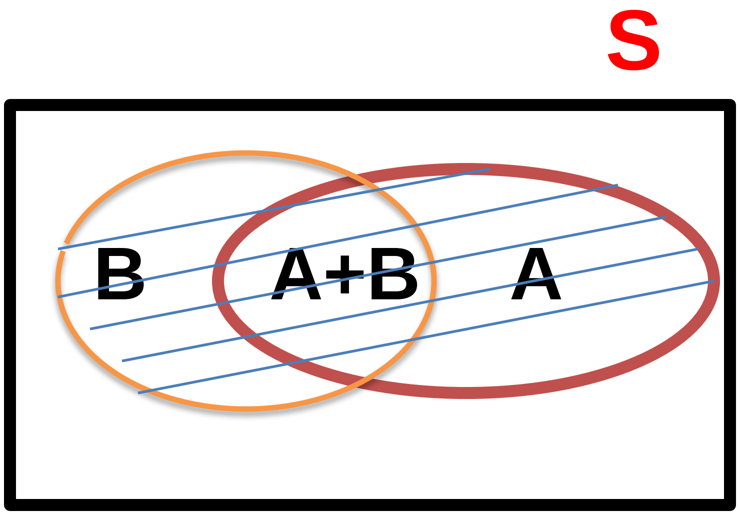
- تركيب الأحداث.
- علاقات الأحداث.
- بعض قوانين الإحتمالات.

تركيب الأحداث

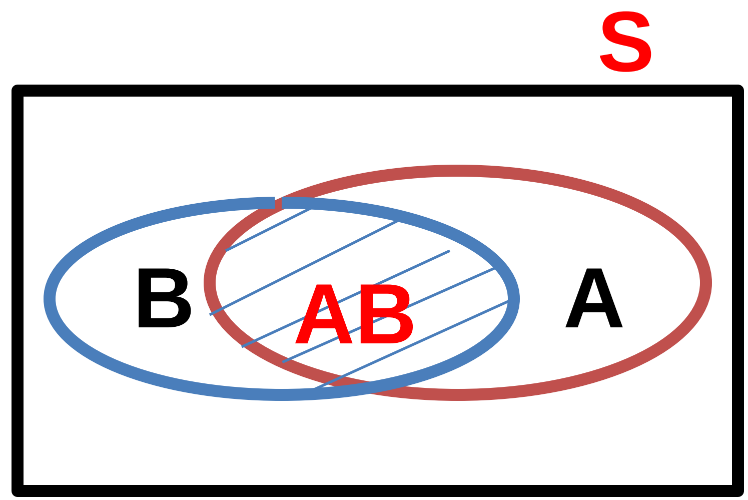
١- في أغلب الأحيان يكون مفيدا أن **نعرض** الحدث بدلالة أحداث أخرى

أ - سيرمز لتقاطع الحدثين AوB بالرمز AB.

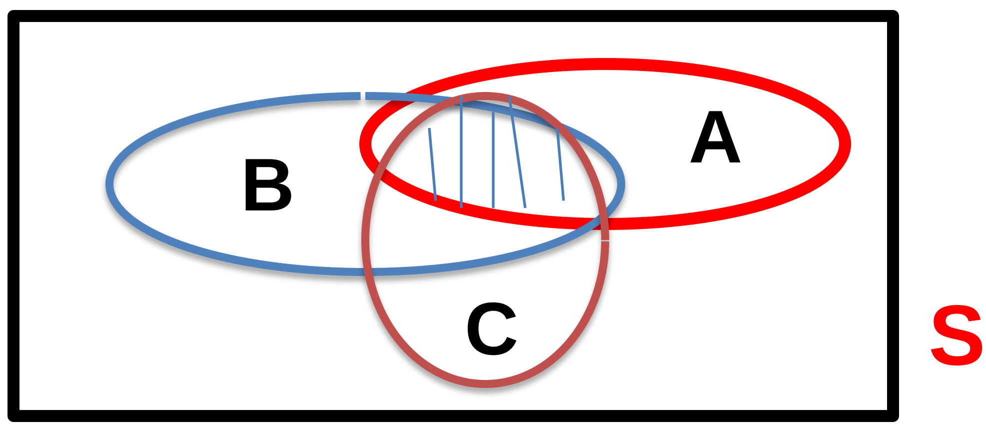
ب- سيرمز لإتحاد الحدثين AوB بالرمز $A+B$.



A+B إتحاد حدثين



AB تقاطع حدثين



ABC

علاقات الأحداث

الأحداث المتكاملة

١- يقال إن الحدث **A** **مكمل** لفئة جميع نقاط العينة التي **لا توجد** في **A** .

2- الرمز $\bar{A}$ يستخدم للدلالة على **مكملة الحدث A** .

3- العلاقة $P(A)+P(\bar{A})=1$ دائما **صحيحة** و علية فإن

$$P(A)=1-P(\bar{A})$$

الأحداث المتنافية

١- يقال إن الحدثين A و B **متنافيان** (غير متقاطعين) إذا كانا **لايحتويان على أي نقاط مشتركة** من نقاط العينة .

٢- إذا كان الحدثان A وB **متنافيان** (غير متقاطعين) فإن الفئة $AB = \varphi$.

٣- بما أن احتمال أي حدث يساوى **مجموع** احتمالات نقاط العينة الذي يحتويها وبالتالي فإنه إذا كان الحدثان AوB **متنافيان** (غير متقاطعين) فإن $P(AB) = 0$.

٤- الحدثان A و $\overline{A}$ **متنافيان** (غير متقاطعين) .

مثال

في تجربة إلقاء عملتين ، إذا كان الحدث A يعنى (على الأقل ظهور صورة واحدة) . أوجد الاحتمال $P(A)$ و $P(\overline{A})$

الحل

أولا - نقاط العينة الأربعة المرتبطين بهذه التجربة هم:

	النتائج	
	العملية رقم 1	العملية رقم 2
E1	H	H
E2	H	T
E3	T	H
E4	T	T

ثانياً - $P(E_i)=\frac{1}{4}, i=1,2,3,4$ (أى ان احتمال أى حدث من الأحداث الأربعة يساوى $\frac{1}{4}$).

ثالثاً - بما أن الحدث A يعنى (على الأقل ظهور صورة واحدة) فإن

$$A = \{E_1,E_2,E_3\}$$

وعلية فإن الحدث $\overline{A}$ تعنى (عدم ظهور أي صورة) ، أي أن

$$\overline{A}=\{E4\}$$

رابعاً - $P(A)=\frac{3}{4}$ و $P(\overline{A})=\frac{1}{4}$.

بعض علاقات (قوانين) الإحتمالات

١- إذا كان $S \subset C$ (C حدث في فراغ العينة S) فإن $P(C) = 1 - P(\overline{C})$

٢- احتمال وقع الحدث المستحيل يساوى صفرا . أي أن ، $P(\Phi) = 0$

٣- إذا كان C_1 وC_2 حدثين من فراغ العينة S بحيث أن $C_1 \subset C_2$ فإنه

$$P(C_1) \leq P(C_2)$$

٤- إذا كان $S \subset C$ (C حدث في فراغ العينة S) فإن $0 \le P(C) \le 1$

٥- إذا كانت C_1 وC_2 حدثين من فراغ العينة S، فإنه ،

$$P(C_1 + C_2) = P(C_1) + P(C_2) - P(C_1 C_2)$$

تمت

بحمد الله

مقدمة في الإحصاء- ريض ٢٠٧
مبادئ الإحصاء- ريض ١٣٠

لكليات العلوم و الاداب

د. عصام الدين احمد ابراهيم رخا

١٤٣٠ – ١٤٣١هـ

الباب الثالث - المحاضرة الثانية عشر

تطبيقات علم الإحصاء (الإحصاءات الحيوية)

عناصر المحاضرة

- مقدمة.
- السكان والتعداد.
- وقت التعداد.
- أساس عمل التعداد.
- البيانات الأساسية في التعداد.

مقدمة

تأتى كلمة **الإحصاءات الحيوية** من كون هذه الإحصاءات تتعلق **بالسكان** وعنصر السكان **حيوي** لعدة أسباب منها أنه:

- ديناميكي أي حركي

، أي أنه

دائم التغير وسريع التغير

أيضا فما من لحظات تمر إلا ويحدث بها إما **ميلاد جديد** أو وفاة حدثت أو **إصابة بمرض معين** أو **حدثت زيجة** أو حدث طلاق.

ـ عنصر السكان هو العنصر الذي يسبب ويحل المشاكل

فما من شك أن الحياة قائمة بالسكان ولهم فالكل يعمل لإسعاد نفسه وغيرة والحكومات كل على شاكلته

يعمل ويخطط

لإرضاء الشعوب ورفع مستواها بين سائر الشعوب الأخرى ، وما من تقدم أو استقرار في الحياة أو تغلب على مشاكل قائمة أو محتملة إلا وكان وراءه

تخطيط دائم

وتنفيذ لهذه التخطيط

ومتابعات مستمرة لذلك التنفيذ

والذي يقوم بكل ذلك هم السكان .

وحتى يقوم التخطيط على **أساس سليم** لابد من توفر **معلومات عن السكان** الذين يخدمهم هذا التخطيط سواء كان هؤلاء السكان **موجودين بالفعل أو محتملين** ، فالتخطيط

للتعليم و للصحة و للمرافق و للخدمات

يحتم توافر

معلومات أساسية عن هؤلاء السكان .

ويتم توفير هذه المعلومات عن السكان وخصائصهم عن طريق ما يسمى

بالتعداد .

السكان والتعداد

المقصود **بالتعداد السكاني** هو

جمع كل البيانات المتعلقة بالخصائص الأساسية للسكان داخل حدود بلد معين في وقت معين

والمقصود **بالخصائص الأساسية** هو

كل ما يتعلق بالفرد من بيانات مثل العمر والنوع والحالة الاجتماعية والحالة الاقتصادية وكذلك معرفة محل الإقامة وعدد أفراد الأسرة ومهنة كل فرد .

وقت التعداد

:

عند القيام بالتعداد يحدد **وقت معين** لإجرائه وعادة ما يكون هذا الوقت

منتصف ليلة معينة في تاريخ محدد

وعادة ما يختار هذا التاريخ من السنة بحيث يكون به

أقل حركة وتنقل ممكن

وبذا **يتواجد معظم السكان في محال إقامتهم العادية**.

وبحيث أنه عند إجراء التعداد يقوم **العدادون** بسؤال المواطنين عن الخصائص **الأساسية** لهم وقت **إجراء** التعداد **وليس** وقت **وصول** **العداد** إلى محل السكن .

أساس عمل التعداد

يوجد أساسان لعمل التعداد هما:

الأساس النظري

و

الأساس الفعلي (العملي)

الأساس النظري

وفيه يتم رد كل فرد إلى محل إقامته عند أخذ البيانات منه .

أي أنه إذا حضر العداد إلى وحدة سكنية ووجد بها إناس غير أهل المسكن فلا يقوم بأخذ أي بيانات منهم

وعلى العكس من ذلك إذا كان أحد ساكني هذه الوحدة غائب لتواجده في مكان أخر غير محل إقامته هذا فيقوم العداد بجمع البيانات المطلوبة عنه ويعد من ساكني هذه الوحدة .

من هذا يتضح أن الأساس النظري يتطلب عدادين مهرة ومدربين جيدا حتى يتمكنوا من رد كل فرد إلى مسكنه أو محل إقامته الأصلي . فإذا لم يتوفر هذا الشرط يلجأ إلى الأساس الثاني للتعداد .

الأساس الفعلي (العملي)

وفيه يتم جمع البيانات الأساسية من الأفراد

المتواجدين في الوحدة السكنية لحظة التعداد بحسب تواجدهم

ولا يلزم رد أي فرد غير متواجد في محل إقامته إلى محل إقامته الأصلي .

وهذا الأساس **لا يحتاج إلى عدادين مهرة** ويتم العمل به في الدول **غير المتقدمة**.

البيانات الأساسية في التعداد

لما كان الهدف من التعداد هو **توفير البيانات الأساسية** عن السكان والتي تحتاجها الدولة في **التخطيط لتوفير الخدمات المناسبة** لهؤلاء السكان وكذلك **عمل المقارنات مع بيانات التعدادات السابقة** حيث أن التعداد **يجرى بصفة دورية عادة كل عشر سنوات** وتلزم أيضا لعمل **المقارنات الدولية** فإن **استمارة التعداد** تشمل على البيانات الأساسية التالية :

- بيانات عامة عن عنوان أفراد الأسرة سواء كانوا يعملون بداخل الدولة أو خارجها .
- الاسم (عادة ثلاثي)
- النوع (ذكر أو أنثى)
- صلة الفرد بعائل الأسرة
-الجنسية
-الديانة
- تاريخ ومحل الميلاد
- السن بالسنوات الكاملة
- مدة الإقامة في محل السكن
- محل الإقامة السابق
- محل الإقامة وقت إجراء التعداد
- الحالة الاجتماعية

- السن عند أول زواج
- عدد الزوجات اللآئى في العصمة
- مدة الحياة الزوجية
- عدد المواليد أحياء بحس النوع
- عدد الباقية منهم على قيد الحياة
- الحالة التعليمية للزوج الحالي أو أخر زوج
- الحالة العملية ومحل العمل ومدة العمل بكل مهنة
- وسيلة الانتقال إلى مقر العمل
- العاهات الظاهرة

-بيانات عن نوع السكن وحيازته والإيجار الشهري وعدد الحجرات والإضاء ومواصفات السكن بالداخل .

تمت

بحمد الله