

## أساسيات الإحصاء

### FUNDAMENTALS OF STATISTICS

#### INTRODUCTION

#### مقدمة

#### Definition of Statistics

#### تعريف الإحصاء

كلمة إحصاء statistics لها معنيان عامان مقبولان :

١- تجمع بيانات كمية خاصة بأى موضوع أو مجموعة، وبصفة خاصة عندما تجمع البيانات وترتب بطريقة منطقية. وأمثلة هذا المعنى هي إحصائيات ضغط الدم، وإحصائيات مباريات كرة القدم، وإحصائيات العمالة، وإحصائيات الحوادث، وما هذه إلا قلة فقط.

٢- العلم الذى يتعامل مع تجميع، وجدولة، وتحليل، وتفسير، وتقديم بيانات كمية.

وقد لوحظ أن المعنى الثانى أوسع عن الأول، حيث أنه يهتم أيضا بتجميع بيانات. ويتعامل استخدام الإحصاء فى مراقبة الجودة مع المعنى الثانى العريض ويشمل أجزاء التجميع، والجدولة، والتحليل، والتفسير، وتقديم البيانات الكمية. ويعتمد كل جزء

على دقة واكتمال الجزء السابق له. فيمكن جمع البيانات بواسطة أحد العاملين في الفحص يقيس مقاومة الشد لأحد الأجزاء البلاستيكية أو بواسطة باحث تسويق يحدد أفضلية العملاء للألوان. ويمكن أن تجردول بأساليب الورقة والقلم البسيطة أو باستخدام الحاسوب. ويمكن أن يشمل التحليل فحصا مرثيا سريعا أو حسابات مرهقة. والنتائج النهائية تفسر وتقدم للمساعدة في اتخاذ قرارات خاصة بالجودة.

وهناك مرحلتان للإحصاء :

١- إحصاء وصفي أو استدلالى descriptive or deductive statistics والذى يهدف إلى وصف وتحليل الموضوع أو المجموعة.

٢- إحصاء استقرائى inductive statistics ،والذى يهدف إلى تحديد، من كمية قليلة نسبيا من البيانات (عينة)، تعليق مهم عن كم أكبر كثيرا من البيانات (المجتمع). وحيث أن هذه التعليقات أو الاستنتاجات لايمكن أخذها بتأكيد مطلق، فعادة ما تستخدم لغة الاحتمالات probability .

ويغطى هذا الفصل أساسيات الإحصاء اللازمة لفهم أساليب مراقبة الجودة التى تليها. وتناقش أساليب الاحتمالات فى الفصل الرابع. وفهم الإحصاء ضرورى لفهم مراقبة الجودة، ولهذا السبب، العديد من المجالات الأخرى.

## Collection of Data

## تجميع البيانات

يمكن جمع البيانات عن طريق الملاحظات المباشرة أو بصورة غير مباشرة عن طريق الأسئلة المكتوبة أو الشفوية. والطريقة الأخيرة تستخدم بكثرة بواسطة الأفراد العاملين فى أبحاث السوق ومستطلعى رأى الجمهور لمعرفة رأى العام. ويتم الحصول على البيانات التى تجمع لأغراض مراقبة الجودة بواسطة الملاحظة المباشرة وتقسّم إما إلى متغيرات أو خواص. والمتغيرات variables هى صفات الجودة التى تقاس، مثل

الوزن الذي يقاس بالجرامات. والخواص attributes، من ناحية أخرى، هي صفات الجودة التي تقسم على أنها إما مطابقة أو ليست مطابقة لمواصفات مثل «قياس المرور/ عدم المرور».

والمتغير القادر بأى درجة على القسمة الجزئية يشار إليه بأنه مستمر continuous . فوزن إحدى مسبوكات حديد الزهر، الذي يمكن أن يقاس بأنه 11 كج، أو 11.33 كج، أو 11.8933 كج (25 رطل)، اعتمادا على دقة آلة الوزن، يكون مثالا لمتغير مستمر. والقياسات مثل الأمتار (الأقدام)، والليترات (الجالونات)، والبسكالات (رطل على البوصة المربعة) هي أمثلة لبيانات مستمرة. والمتغيرات التي تعرض فجوات تسمى وثابة discrete . عدد مسامير البرشام غير المطابقة في مقطورة سفر يمكن أن يكون أى رقم، مثل 0 أو 3 أو 5 أو 10 أو 96 أو غيرها، إلا أنه لا يمكن أن يكون هناك 4.65 مسمار برشام ليس مطابق في مقطورة معينة. وعموما البيانات المستمرة تكون قابلة للقياس، بينما الوثابة تكون قابلة للعد.

وفي بعض الأحيان يكون مقنعا للبيانات الشفوية أو غير العددية أن تفترض طبيعة المتغير. مثال ذلك، جودة السطح النهائي لقطعة من الاثاث يمكن أن تقسم على أنها ضعيفة، أو متوسطة، أو جيدة. وتقسيم الضعيف، والمتوسط، والجيد يمكن أن يستبدل بقيم عددية مثل 1 و 2 و 3 على التوالي. وبطريقة شبيهة، تشير المنشآت التعليمية إلى رموز المعدلات A، B، C، و D، و E، بالقيم العددية، 0، 1، 2، 3، 4 على التوالي، وتستخدم هذه القيم العددية الوثابة كمتغيرات وثابة فى أغراض الحسابات.

وبينما يعبر عن العديد من صفات الجودة بواسطة متغيرات، إلا أنه هناك العديد من الصفات التي يجب التعبير عنها كخواص. وتكراريا، هذه الصفات التي يحكم عليها بالملاحظة البصرية تقسم على أنها خواص. فالسلك على موتور كهربائى إما أن

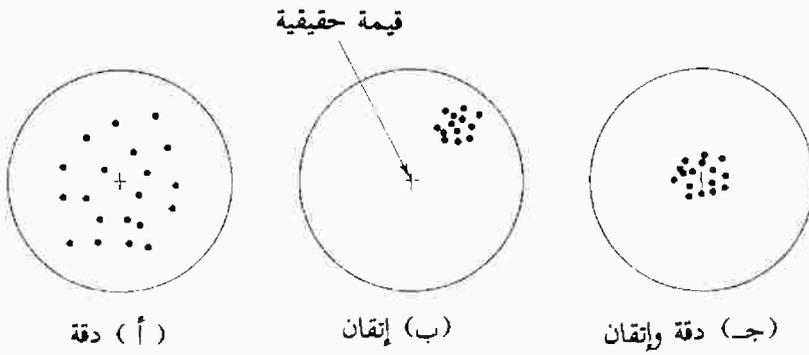
يكون متصلًا بنهاية طرفية أو غير متصل، والكلمات في هذه الصفحة إما أن تكون مكتوبة بهجاء صحيح أو غير صحيح، والمفتاح الترددي إما أن يكون في وضع التوصيل أو غير التوصيل، والإجابة إما أن تكون صحيحة أو خطأ. والأمثلة المعطاه في الجملة السابقة تبين مطابقة لمواصفة معينة وعدم مطابقة لنفس المواصفة.

ويكون من المرغوب فيه في بعض الأحيان أن تقسم المتغيرات كخواص. ويتكرر أن يكون العاملون في المصنع مهتمون بمعرفة ما إذا كان المنتج الذين ينتجونه مطابقا للمواصفات أم لا. مثال ذلك، القيمة العددية لوزن لفة من السكر قد لا تكون بنفس أهمية المعلومات بأن الوزن يقع داخل حدود سبق تحديدها. لهذا، البيانات، التي تجتمع عن وزن لفة السكر، يذكر أنها مطابقة أو ليست مطابقة للمواصفات.

في تجميع البيانات يكون عدد الأرقام دالة للاستخدام المطلوب للبيانات. مثال ذلك، عند تجميع بيانات عن عمر اللمبات الكهربائية المنزلية، يكون مقبولا تسجيل 995.6 ساعة، إلا أن تسجيل قيمة 995.632 ساعة يكون أكثر دقة وغير ضروري. بالمثل، إذا كانت إحدى المواصفات الرئيسية لها حد أدنى 9.52 م (0.375 بوصة) وحد أعلى 9.85 م (0.377 بوصة)، فيمكن جمع بيانات لأقرب 0.001 م والتقريب إلى أقرب 0.01 م. وعموما، كلما ازدادت الأرقام على يمين العلامة العشرية، كلما تعقد جهاز القياس.

وقد لا تعطى أجهزة القياس قراءة حقيقية بسبب مشاكل خاصة بالدقة والإتقان. ويبين شكل ٢ - ١ - أ سلسلة دقيقة من القياسات المتكررة بسبب أن متوسطها يكون قريبا من القيمة الحقيقية، والتي تكون في المركز. وفي شكل ٢ - ١ - ب تكون القياسات المتكررة في السلسلة قريبة جدا من بعضها إلا أنها ليست قريبة من القيمة الحقيقية. ويبين شكل ٢ - ١ - ج سلسلة القياسات المتكررة مضغوطة بشدة حول القيمة الحقيقية وهذه القياسات تكون دقيقة ومتقنة.

وتقريب البيانات يتطلب اتباع تحويل معين. ففي تقريب الأعداد 0.9530 و0.9531 و0.9532 و0.9533 و0.9534 إلى أقرب واحد من ألف، تكون النتيجة 0.953 حيث أن كل الأعداد أقرب إلى 0.953 عنها إلى 0.954 . وعند تقريب الأعداد 0.9535 و0.9536 و0.9537 و0.9538 و0.9539 تكون النتيجة 0.954 ، حيث أن كل الأعداد أقرب إلى 0.954 عنها إلى 0.953 . وفي كلمات أخرى، إذا كان الرقم التالي 5 أو أكبر، فيقرب العدد لأعلى.



شكل ١.٢: الفرق بين الدقة والإتقان

عند العمل ببيانات عددية، تكون الأرقام المعنوية significant figures مهمة جدا. والأرقام المعنوية لعدد هي الأرقام مع استثناء أى أصفار رائدة تلزم لتحديد موقع العلامة العشرية. مثال ذلك العدد 3.69 له ثلاثة أرقام معنوية، والعدد 36.900 له خمسة أرقام معنوية، والعدد 2700 له أربعة أرقام معنوية، والعدد 22.0365 له ستة أرقام معنوية، والعدد 0.00270 له ثلاثة أرقام معنوية. الأصفار فى نهاية العدد تحسب على أنها أرقام معنوية، بينما الأصفار الرائدة لا تحسب كأرقام معنوية. وتقدم هذه القاعدة بعض الصعوبة عند العمل مع أرقام صحيحة حيث أن العدد 300 يكون له رقم معنوى واحد أو اثنين أو ثلاثة. ويمكن إلغاء هذه الصعوبة باستخدام التمثيل العلمى. له فإن

$3 \times 10^2$  يكون له رقم معنوى واحد، والعدد  $3.0 \times 10^2$  يكون له رقمان معنويان، والعدد  $3.00 \times 10^2$  يكون له ثلاثة أرقام معنوية. والأعداد التي تصاحب العد يكون لها عدد لانهاى من الأرقام المعنوية، وحتى العدد 65 يمكن أن يكتب على أنه 65.0000 بأى عدد أصفار، لانهاى، على يمين العلامة العشرية.

وعند تنفيذ العمليات الرياضية للضرب، والقسمة، والعمليات الاسية، يكون للإجابة نفس عدد الأرقام المعنوية مثل أقل أرقام معنوية مستخدمة. والأمثلة التالية تساعد فى توضيح هذه القاعدة.

$$\sqrt{81.9} = 9.05$$

$$6.59 \times 2.3 = 15$$

$$(24 \text{ ليس عدد للعد}) \quad 32.65 \div 24 = 1.4$$

$$(24 \text{ هو عدد للعد له القيمة } 24.00000) \quad 32.65 \div 24 = 1.360$$

عند تنفيذ العملية الرياضية للجمع والطرح، فيمكن ألا يكون للإجابة النهائية أكثر من أرقام معنوية بعد العلامة العشرية عن العدد الذى له أقل أرقام معنوية بعد العلامة العشرية، وفى الحالات التى تشمل أعدادا بدون علامات عشرية، لا يكون للإجابة النهائية أرقام معنوية أكثر من العدد الذى له أقل عدد للأرقام المعنوية والأمثلة لتوضيح هذه القاعدة هى ما يلى :

$$(6 \text{ ليس عدد للعد}) \quad 38.26 - 6 = 32$$

$$(6 \text{ عدد للعد}) \quad 38.26 - 6 = 32.26$$

$$(تم تقريب الإجابة من 32.16) \quad 38.26 - 6.1 = 32.2$$

$$(أقل أرقام معنوية هى 2) \quad 8.1 \times 10^3 - 1232 = 6.9 \times 10^3$$

$$(أقل أرقام معنوية هى 4) \quad 8.100 \times 10^3 - 1232 = 6868$$

استخدام القواعد السابقة يمنع التعارض في الإجابات عند أفراد مراقبة الجودة، إلا أن بعض الحكم يمكن أن يكون لازماً في بعض الأحيان. وفي أى حالة، يمكن ألا تكون الإجابة أكثر دقة عن البيانات القادمة.

## Describing the Data

## وصف البيانات

في الصناعة، والأعمال، والمصالح الحكومية كميات البيانات التي جمعت لها أحجام هائلة. حتى عنصر واحد، مثل عدد أجراس الدخان لعدم المطابقة في 35 دفعة بكل منها 1000، يمكن أن يمثل مثل هذا الكم الكبير من البيانات الذي يمكن أن يكون مثيراً للخلط أكثر من كونه مفيداً. اعتبر البيانات المبينة في جدول ٢ - ١. من الواضح، أن هذه البيانات، في هذه الصورة، تكون صعبة الاستخدام ولا تكون فعالة في وصف خواص البيانات. وتلزم بعض وسائل تلخيص البيانات لتوضيح أى قيمة أو قيم تتجمع البيانات حولها أو تشتتت أو تبتعد البيانات عنها. وهناك أسلوبان متاحان لتحقيق هذا التلخيص والرسومات والتحليل.

جدول ٢ - ١ : عدد أجراس الدخان لعدم المطابقة في 35 دفعة بكل منها 1000

|   |   |   |   |   |
|---|---|---|---|---|
| 0 | 0 | 3 | 1 | 0 |
| 2 | 1 | 4 | 5 | 0 |
| 0 | 0 | 2 | 0 | 1 |
| 2 | 1 | 1 | 1 | 2 |
| 1 | 3 | 0 | 4 | 0 |
| 2 | 1 | 0 | 3 | 1 |
| 0 | 3 | 0 | 4 | 0 |

أسلوب الرسم عبارة عن رسم أو تصوير التوزيع التكرارى frequency distribution ، وهو ملخص لكيفية حدوث نقاط (ملاحظات) البيانات داخل كل جزء فرعى لنقيم التى تمت ملاحظتها أو لمجموعات القيم التى تمت ملاحظتها. ويلخص الأسلوب التحليلى البيانات عن طريق حساب قياس النزعة المركزية -measure of central tendency وقياس التشتت measure of dispersion. وفى بعض الأحيان تستخدم الوسيلتان، وسيلة الرسم والوسيلة التحليلية.

وتوصف هاتان الوسيلتان فى الأقسام التالية من هذا الفصل.

## FREQUENCY DISTRIBUTION

## التوزيع التكرارى

### Ungrouped Data

### البيانات غير المجمعة

تشمل البيانات غير المجمعة قائمة بالقيم التى تمت ملاحظتها، بينما تمثل البيانات المجمعة تجميعا للبيانات التى تمت ملاحظتها. ويمكن أن تكون البيانات وثابة، كما هو الحال فى هذا القسم، أو مستمرة، كما هو الحال فى القسم التالى.

ونظرا لأن البيانات غير المنظمة لامعنى لها افتراضيا، فيلزم طريقة لتشغيل البيانات. ويستخدم جدول ٢ - ١ فى توضيح المفهوم. فالخلل الذى يراجع المعلومات كما هى معطاه فى هذا الجدول يكون لديه صعوبة فى فهم معنى البيانات. ويمكن الحصول على فهم أفضل كثيرا بعد عدد مرات تكرار كل قيمة، كما هو مبين فى جدول ٢ - ٢.

جدول ٢.٢: عد عدد أجراس الدخان غير المطابقة

| عدد وحدات | الجدولة | التكرار |
|-----------|---------|---------|
| 0         |         | 13      |
| 1         |         | 9       |
| 2         |         | 5       |
| 3         |         | 4       |
| 4         |         | 3       |
| 5         |         | 1       |

الخطوة الأولى هي عمل منظومة array، وهي ترتيب لبيانات عددية خام ترتيباً تصاعدياً أو تنازلياً طبقاً لقيمتها. المنظومة التي لها ترتيب تصاعدي من 0 إلى 5 مبيّنة في العمود الأول من جدول ٢ - ٢.

الخطوة التالية هي جدولة تكرار كل قيمة عن طريق وضع علامات للعد تحت عمود الجدولة وفي الصف المناسب من جدول ٢ - ٢. ابدأ بالأعداد 0 و 0 و 1 و 2 و... من جدول ٢ - ١ واستمر في وضع علامات العد حتى تتم جدولة البيانات. آخر عمود في جدول ٢ - ٢ هو قيمة عددية لعدد العدادات ويسمى التكرار frequency.

ويبين تحليل جدول ٢ - ٢ أنه يمكن رؤية توزيع البيانات. فإذا حذف عمود «الجدولة»، فيقسم الجدول الناتج على أنه توزيع تكراري frequency distribution، وهو عبارة عن ترتيب للبيانات بين تكرار القيم في كل فئة.

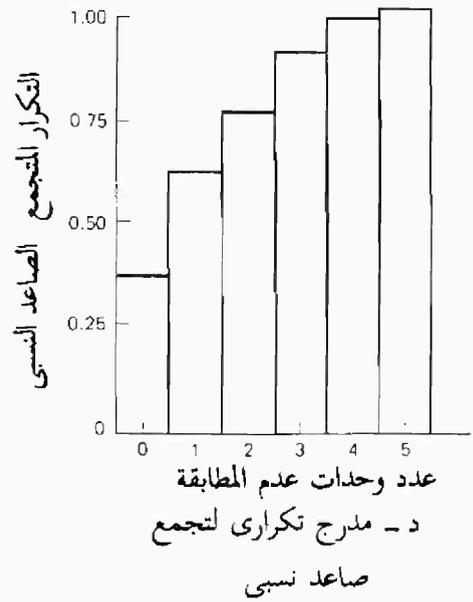
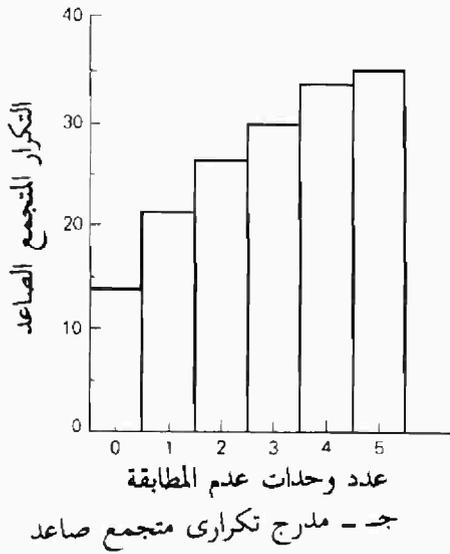
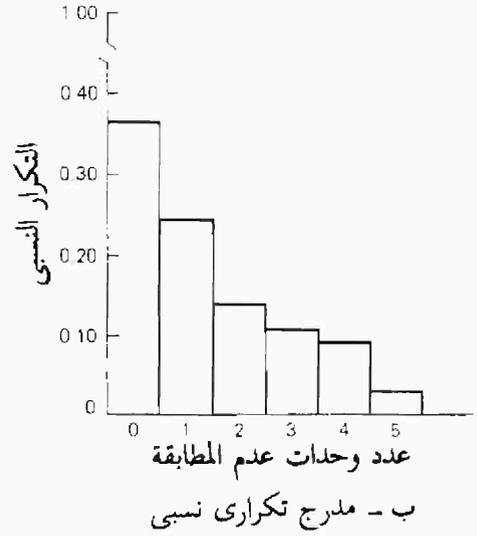
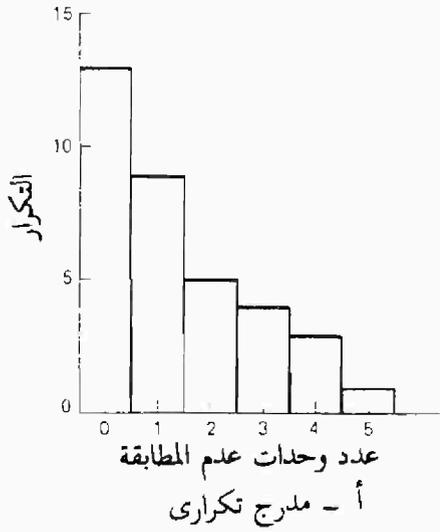
والتوزيع التكراري هو طريقة مفيدة لرؤية البيانات وهو مفهوم إحصائي أساسي. التفكير في فئة إعداد يكون لها نوع معين من التوزيع يكون أساسياً في حل مشاكل مراقبة الجودة. فهناك أنواع مختلفة للتوزيعات التكرارية، ويمكن أن يحدد نوع التوزيع التكراري طريقة حل المشكلة.

وتقدم التوزيعات التكرارية في صورة رسومات عندما يكون هناك رغبة في توضيح مرئى أكبر. وهناك عدد من الطرق المختلفة لتقديم التوزيعات التكرارية.

المدرج التكرارى histogram يحتوى على مجموعة من المستطيلات التى تمثل التكرار فى كل فئة. وهو يمثل بالرسم التكرارات للقيم التى تمت ملاحظتها. وشكل ٢ - ٢ - أ هو مدرج تكرارى لبيانات جدول ٢ - ٢. وحيث أن هذا متغير وثاب، فيكون الخط الرأسى فى موقع المستطيل صحيحا نظريا (انظر شكل ٢ - ٥). إلا أن المستطيلات شائعة الاستخدام.

نوع آخر من تمثيل الرسومات هو التوزيع التكرارى النسبى. نسبى، بمعنى تناسب أو كسر من الكل. وتحسب التكرارات النسبية عن طريق قسمة تكرار كل قيمة بيانات (فى هذه الحالة، عدد عدم المطابقة) على الكل، وهو مجموع التكرارات لكل قيمة بيانات. وتظهر الحسابات فى العمود الثالث من شكل ٢ - ٣. والتمثيل بالرسم مبين فى شكل ٢ - ٢ - ب. وللتكرار النسبى ميزة الإشارة. مثال ذلك، تناسب أن وحدتى عدم مطابقة يكون 0.14. ويفضل بعض الممارسون استخدام نسب مئوية للمقياس الراسى بدلا من الكسور.

التكرار المتجمع الصاعد يحسب عن طريق جمع تكرار كل قيمة بيانات على مجموع التكرارات لقيم البيانات السابقة. وكما هو مبين فى العمود الرابع من جدول ٢ - ٣، التكرار المتجمع الصاعد لوحدها عدم المطابقة التى عددها 0 هو 13، والثى عددها 1 هو  $22 + 9 = 31$  والثى عددها 2 هو  $22 + 5 = 27$ ، وهكذا. والتكرار المتجمع الصاعد هو عدد نقاط البيانات الذى يساوى أو يكون أقل من قيمة البيانات. مثال ذلك، عدد الدفعات التى لها 2 أو أقل من وحدات عدم المطابقة هو 72. وتمثيل الرسومات مبين فى شكل ٢ - ٢ - ج.



شكل ٢٠٢: تمثيل بالرسومات لبيانات معطاه فى جدولى ٢٠٢ و ٢٠٢

جدول ٣٠٢ : توزيعات تكرارية مختلفة لبيانات معطاه في جدول ٢ - ١

| عدد وحدات<br>عدم المطابقة | التكرار | التكرار النسبي      | التكرار المتجمع<br>الصاعد | التكرار المتجمع<br>النسبي |
|---------------------------|---------|---------------------|---------------------------|---------------------------|
| 0                         | 13      | $13 \div 35 = 0.37$ | 13                        | $13 \div 35 = 0.37$       |
| 1                         | 9       | $9 \div 35 = 0.26$  | $13 + 9 = 22$             | $22 \div 35 = 0.63$       |
| 2                         | 5       | $5 \div 35 = 0.14$  | $22 + 5 = 27$             | $27 \div 35 = 0.77$       |
| 3                         | 4       | $4 \div 35 = 0.11$  | $27 + 4 = 31$             | $31 \div 35 = 0.89$       |
| 4                         | 3       | $3 \div 35 = 0.09$  | $31 + 3 = 34$             | $34 \div 35 = 0.97$       |
| 5                         | 1       | $1 \div 35 = 0.03$  | $34 + 1 = 35$             | $35 \div 35 = 1.00$       |
| الإجمالي                  | 35      | 1.00                |                           |                           |

ويحسب التكرار النسبي المتجمع الصاعد عن طريق قسمة التكرار المتجمع الصاعد لكل قيمة بيانات على الإجمالي. هذه الحسابات مبينة في العمود الخامس من جدول ٢ - ٣، وتمثيل الرسوم مبين في شكل ٢ - ٢ - د. ويبين الرسم أن تناسب دفعات أجراس الدخان التي لها وحدتا عدم مطابقة أو أقل يكون 0.77 أو 77%.

المثال السابق يكون محدودا بمتغير وثاب له ست قيم. وبالرغم من أن هذا المثال يكون كافيا كمقدمة أساسية لمفهوم التوزيع التكراري، إلا أنه لا يقدم معرفة فاحصة بالموضوع.

### Grouped Data

### البيانات المجمعة

تكوين التوزيع المتجمع الصاعد للبيانات المجمعة يكون أكثر تعقيدا نظرا لأنه يوجد في العادة عدد أكبر من الفئات. ويوضح مثال لمشكلة تستخدم متغير مستمر المفهوم.

١- إجمع البيانات وقم بإعداد ورقة بالعدد collect data and construct a tally sheet. البيانات المجموعة عن أوزان ١١٠ قضيب صلب مبنية في جدول ٢ - ٤ . الخطوة الأولى هي أعداد عدد القيم، كما هو مبين في جدول ٢ - ٥ . ولكي نكون أكثر فاعلية، تشفر الأوزان من 2.500 Kg كج، وهذه طريقة مستخدمة في تبسيط البيانات. لهذا، فالوزن الذي قيمته 31 يكون مكافئاً لوزن 2.531. كج + 2.500 (0.031). ويبين تحليل جدول ٢ - ٥ أنه تنقل معلومات أكثر إلى المحلل عما تنقل من بيانات جدول ٢ - ٤، إلا أن الصورة لاتزال غير واضحة بعض الشيء.

في هذه المشكلة يوجد 45 فئة، وهي كثيرة جداً ويجب أن تقلل عن طريق عمل التجميع في خلايا<sup>(١)</sup> والخلية هي تجميع داخل حدود محددة للقيم الملاحظة عبر المحور الأفقى للمدرج التكرارى (انظر شكل ٢ - ٣). وتجميع البيانات في خلايا يبسط تمثيل التوزيع، إلا أن بعض التفاصيل تفقد. وعندما يكون عدد الخلايا كبيراً، تنحرف الصورة الحقيقية للتوزيع عن طريق الخلايا التي يكون لها عدد غير كاف من العناصر أو لا يكون بها عناصر على الإطلاق. أو، عندما يكون عدد الخلايا صغيراً، يتركز عدد كبير جداً من العناصر في قلة من الخلايا وينحرف التوزيع أيضاً.

عدد الخلايا أو المجموعات في توزيع تكرارى يكون نوعاً من الحكم الشخصى بصورة كبيرة من المحلل. هذا الحكم يكون مبنياً على عدد الملاحظات ويمكن أن يحتاج إلى المحاولة والخطأ لتحديد العدد الأمثل للخلايا. وعموماً، يجب أن يقع عدد الخلايا بين 5 و 20 . والخطوط الإرشادية العريضة هي كما يلي : يستخدم من 5 إلى 9 خلايا عندما يكون عدد الملاحظات أقل من 100 ، ويستخدم من 8 إلى 17 خلية عندما يقع عدد الملاحظات بين 100 و 500 ، ويستخدم من 15 إلى 20 خلية عندما يكون عدد الملاحظات أكبر من 500 . ولتوفير المرونة، يتداخل عدد الخلايا

(١) تستخدم كلمة فئة class في بعض الأحيان بدلا من كلمة خلية cell

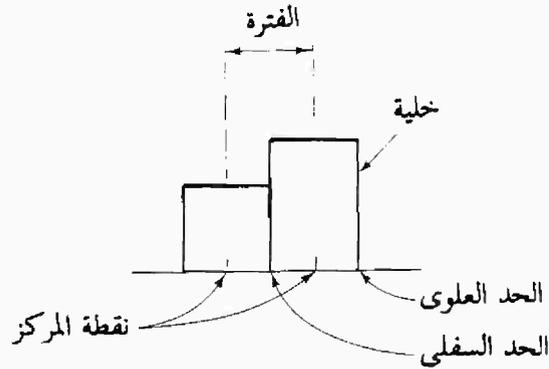
في الخطوط الإرشادية. وهذا يعني أن هذه الخطوط الإرشادية ليست متحركة ويمكن أن تضبط عند الحاجة لتقديم توزيع تكرارى مقبول.

جدول ٤.٢: أوزان قضبان الصلب (بالكيلوجرام)

|       |           |           |       |       |
|-------|-----------|-----------|-------|-------|
| 2.559 | 2.556     | 2.566     | 2.546 | 2.561 |
| 2.570 | 2.546     | 2.565     | 2.543 | 2.538 |
| 2.560 | 2.560     | 2.545     | 2.551 | 2.568 |
| 2.546 | 2.555     | 2.551     | 2.554 | 2.574 |
| 2.568 | 2.572     | 2.550     | 2.556 | 2.551 |
| 2.561 | 2.560     | 2.564     | 2.567 | 2.560 |
| 2.551 | 2.562     | 2.542     | 2.549 | 2.561 |
| 2.556 | 2.550     | 2.561     | 2.558 | 2.556 |
| 2.559 | 2.557     | 2.532     | 2.575 | 2.551 |
| 2.550 | 2.559     | 2.565     | 2.552 | 2.560 |
| 2.534 | 2.547     | 2.569     | 2.559 | 2.549 |
| 2.544 | 2.550     | 2.552     | 2.536 | 2.570 |
| 2.564 | 2.553     | 2.558     | 2.538 | 2.564 |
| 2.552 | 2.543     | 2.562     | 2.571 | 2.553 |
| 2.539 | 2.569     | 2.552     | 2.536 | 2.537 |
| 2.532 | 2.552     | 2.575 (H) | 2.545 | 2.551 |
| 2.547 | 2.537     | 2.547     | 2.533 | 2.538 |
| 2.571 | 2.545     | 2.545     | 2.556 | 2.543 |
| 2.551 | 2.569     | 2.559     | 2.534 | 2.561 |
| 2.567 | 2.572     | 2.558     | 2.542 | 2.574 |
| 2.570 | 2.542     | 2.552     | 2.551 | 2.553 |
| 2.546 | 2.531 (L) | 2.563     | 2.554 | 2.544 |

جدول ٥.٢: صفحة أوزان القضبان الصلب (مشفرة من 2.500 Kg)

| الوزن | الجدولة | الوزن | الجدولة | الوزن | الجدولة |
|-------|---------|-------|---------|-------|---------|
| 31    | I       | 46    | III     | 61    | IV      |
| 32    | II      | 47    | III     | 62    | II      |
| 33    | I       | 48    |         | 63    | I       |
| 34    | II      | 49    | II      | 64    | III     |
| 35    |         | 50    | III     | 65    | II      |
| 36    | II      | 51    | IV III  | 66    | I       |
| 37    | II      | 52    | IV I    | 67    | II      |
| 38    | III     | 53    | III     | 68    | II      |
| 39    | I       | 54    | II      | 69    | III     |
| 40    |         | 55    | I       | 70    | III     |
| 41    |         | 56    | IV      | 71    | II      |
| 42    | III     | 57    | I       | 72    | II      |
| 43    | III     | 58    | III     | 73    |         |
| 44    | II      | 59    | IV      | 74    | II      |
| 45    | III     | 60    | IV      | 75    | II      |



شكل ٣.٢ : تسمية الخلية

٢- حدد المدى  $R$  . determine the range . هو الفرق بين أعلى قيمة ملاحظة وأقل قيمة ملاحظة كما هو مبين بالعلاقة :

$$R = X_H - X_L$$

حيث :  $R$  = المدى

$$X_H = \text{أعلى عدد}$$

$$X_L = \text{أقل عدد}$$

من جدول ٢ - ٤ أو جدول ٢ - ٥ أعلى عدد هو 2.5751 وأقل عدد هو 2.531 . لهذا،

$$\begin{aligned} R &= X_H - X_L \\ &= 2.575 - 2.531 \\ &= 0.044 \end{aligned}$$

٣- حدد فترة الخلية cell interval . فترة الخلية cell interval هي المسافة بين مركزي خليتين متجاورتين كما هو مبين في شكل ٢ - ٣ . وكلما كان ممكننا، فالفترة الفردية مثل 0.001 أو 0.07 أو 0.5 أو 3 يوصى بها بحيث إن قيم المركز تكون لنفس عدد المواقع العشرية مثل قيم البيانات. فترة الخلية (i) وعدد الخلايا (h) يرتبطان ببعض بالعلاقة  $h = R/i$ . حيث كل من h و i غير معروفة، وتستخدم طريقة المحاولة والخطأ في إيجاد الفترة التي تتفق مع الخطوات الإرشادية.

$$h = \frac{R}{i} = \frac{0.044}{0.003} = 15 \quad \text{افرض أن } i = 0.003 \text{ فتكون}$$

$$h = \frac{R}{i} = \frac{0.044}{0.005} = 9 \quad \text{افرض أن } i = 0.005 \text{ فتكون}$$

$$h = \frac{R}{i} = \frac{0.044}{0.007} = 6 \quad \text{افرض أن } i = 0.007 \text{ فتكون}$$

فترة الخلية التي قيمتها 0.005 مع تسع خلايا تعطى أفضل تقديم للبيانات مبنيا على الخطوات الإرشادية لعدد الخلايا المعطى في الخطوة الأولى.

طريقة أخرى للحصول على فترة الخلية هي استخدام قاعدة ستورجى Sturgie ، وهي :

$$i = \frac{R}{1 + 3.322 \log n}$$

وبالنسبة إلى مثال المشكلة تكون النتيجة كما يلي :

$$i = \frac{R}{1 + 3.322 \log n} = \frac{0.044}{1 + 3.322 (2.041)} = 0.0057$$

وتكون أقرب فترة فردية للبيانات هي 0.005 . وتعطى كل من الطريقتين نتائج متشابهة.

٤- حدد مراكز الخلايا determine the cell midpoints . أقل مركز خلية يجب أن يحدد ليشمّل أقل قيمة بيانات في خليته. وأبسط طريقة هي اختيار أقل نقطة

بيانات (2.531) كقيمة مركز لأول خلية. ويوصى بهذه الطريقة للقراء المبتدئين لتوهم عملهم في مجال الإحصاء. وحيث أن الفترة هي 0.005 ، فتوجد خمس قيم بيانات في كل خلية، لهذا، فإن قيمة المركز 2.533 يمكن أن تستخدم لأول خلية. هذه القيمة يكون لها أقل قيمة بيانات (2.531) في أول خلية، والتي يكون لها قيم البيانات 2.531 و 2.532 و 2.533 و 2.534 و 2.535 .

اختيار المركز ما هو إلا موضوع يعتمد على الحكم الخاص، وفي هذه الحالة تم اختيار المركز 2.533 بحيث أن عدد الخلايا يكون 9 . واختيار أى مركز آخر، بالرغم من عدم صحته، سوف يعطى عشر خلايا في التوزيع التكرارى. واختيار قيم مراكز مختلفة ينتج عنه توزيعات تكرارية مختلفة - وهنا خمسة توزيعات ممكنة. ومراكز الخلايا الثمان المتبقية يتم الحصول عليها عن طريق جمع فترة الخلية على المركز السابق :

$$2.533 + 0.005 = 2.538 \text{ و } 2.533 + 0.005 = 2.543 \text{ و } 2.538 + 0.005 = 2.543$$

$$2.548 = \dots \text{ و } 2.573 = 2.568 + 0.005 . \text{ هذه المراكز مبينة في جدول ٢ - ٦ .}$$

**جدول ٦.٢: التوزيع التكرارى لأوزان قضبان الصلب**

| حدود الخلية | نقطة مركز الخلية | التكرار |
|-------------|------------------|---------|
| 2.531-2.535 | 2.533            | 6       |
| 2.536-2.540 | 2.538            | 8       |
| 2.541-2.545 | 2.543            | 12      |
| 2.546-2.550 | 2.548            | 13      |
| 2.551-2.555 | 2.553            | 20      |
| 2.556-2.560 | 2.558            | 19      |
| 2.561-2.565 | 2.563            | 13      |
| 2.566-2.570 | 2.568            | 11      |
| 2.571-2.575 | 2.573            | 8       |
| Total       |                  | 110     |



ويفضل بعض المحللين أن يتركوا الحدود عند نفس عدد المواقع العشرية مثل البيانات. ولا تقابل أى صعوبة مع هذه العملية طالما أن فترة الخلية تكون فردية ويكون مفهوماً أن الحدود الحقيقية تكون متسعة بمنتصف الطريق إلى العدد التالى. وتتبع هذه العملية فى هذا الكتاب. لهذا، فإن الحد السفلى لأول خلية يكون 2.531.

وبمجرد تحديد الحدود لخلية واحدة، يتم الحصول على حدود الخلايا الأخرى بالجمع المتتالى لفترة الخلية. لهذا، تكون الحدود السفلى كما يلي :

$2.536 = 2.531 + 0.005$  و  $2.541 = 2.536 + 0.005$  و ... و  $2.566 = 2.536 + 0.005$  . ويتم الحصول على الحدود العليا بطريقة شبيهة وهى مبينة فى العمود الأول من جدول ٢ - ٦ .

٦- ضع تكرار الخلية *post the cell frequency* . كمية الأعداد فى كل خلية توضع فى عمود التكرار من جدول ٢ - ٦ . وبين تحليل جدول ٢ - ٥ أنه لأقل خلية يكون هناك : قيمة واحدة من 2.531 واثنتان من 2.532 ، وواحدة من 2.533 ، واثنتان من 2.534 ، وصفر من 2.535 . لهذا، يوجد إجمالى ست قيم فى أقل خلية، ويكون للخلية التى لها مركز 2.533 تكرارا مقداره 6 . وتتحدد القيمة لبقية الخلايا الأخرى بطريقة شبيهة.

والتوزيع التكرارى الكامل مبين فى جدول ٢ - ٦ . ويعطى التوزيع التكرارى هذا إدراكا أفضل للقيمة المركزية وكيفية تشتت البيانات حول هذه القيمة عن البيانات غير المنظمة أو صفحة العد. وبين شكل ٢ - ٤ المدرج التكرارى .

والمعلومات عن إعداد المدرجات التكرارية للتكرار النسبى، والتكرار المتجمع الصاعد، والتكرار النسبى التراكمى للبيانات المجمعة تكون نفسها مثل البيانات غير المجمعة مع استثناء واحد. مع المدرجين التكرارين للتكرار المتجمع الصاعد، يكون

الحد العلوى الحقيقى للخلية يكون القيمة المسماة على المحور السينى . وإعداد هذه المدرجات التكرارية لمثال المشكلة متروك للقارئ كتمرين .

ويصف المدرج التكرارى الاختلاف فى العملية ويستخدم فى :

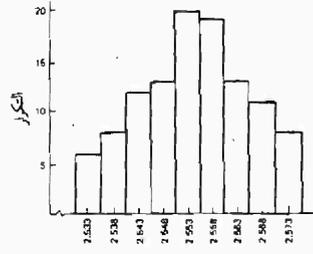
- ١- تحديد مقدرة العملية .
- ٢- المقارنة مع المواصفات .
- ٣- اقتراح شكل المجتمع .
- ٤- تحديد التعارضات فى البيانات مثل الفجوات .

## أنواع أخرى لرسومات التوزيع التكرارى

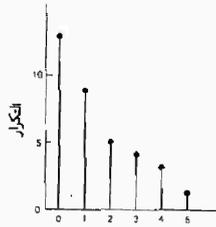
### Other Types of Frequency Distribution Graphs

يمكن أن تمثل خريطة الأعمدة التوزيعات التكرارية أيضا، كما هو مبين فى شكل ٢ - ٥ - أ باستخدام بيانات جدول ٢ - ١ . وكما سبق ذكره، فإن خريطة الأعمدة تكون صحيحة نظريا للبيانات الوثابة إلا أنها ليست شائعة الاستخدام .

المضلع polygon أو المضلع التكرارى frequency polygon هو طريقة رسم أخرى لتقديم التوزيعات التكرارية وهو موضح فى شكل ٢ - ٥ - ب مستخدما بيانات جدول ٢ - ٦ . ويعد عن طريق وضع نقطة على مركز كل خلية عند الارتفاع الذى يحدد كل تكرار . ويمتد المنحنى عند كل نهاية وذلك حتى يصبح الشكل مغلقا . وحيث أن المدرجات التكرارية تبين المساحة فى كل خلية، فتعتبر أنها تقدم صورة مرسومة أفضل عن المضلع التكرارى ولهذا فهى الأكثر استخداما .

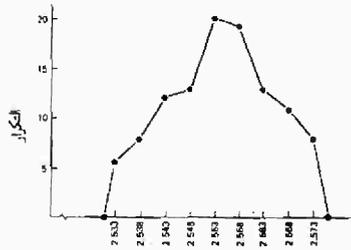


أوزان قضبان الصلب



عدد غير المطابق

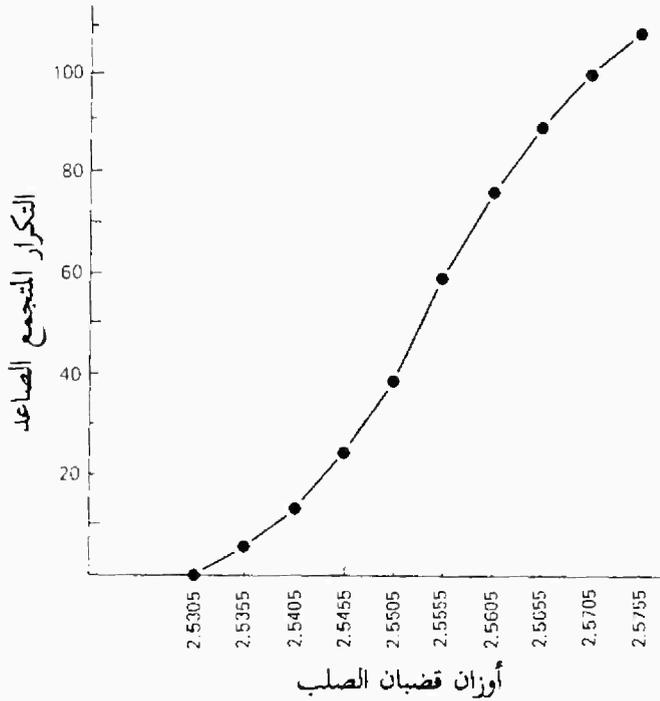
رسم أعمدة بيانات جدول ٢ - ١



أوزان قضبان الصلب

مضلع بيانات جدول ٦٠٢

شكل ٥٠٢: أنواع أخرى لرسومات التوزيع التكراري



(ج) توزيع تكرارى متجمع صاعد لبيانات جدول ٢ - ٦

الرسم المستخدم فى تقديم التكرار لكل القيم الأقل من حد الخلية العلوى لخلية معينة يعرف بأنه تكرار متجمع صاعد cumulative frequency أو أنه منحنى متجمع صاعد ogive . ويبين شكل ٢ - ٥ - ج منحنى توزيع تكرارى متجمع صاعد لبيانات جدول ٢ - ٦ . القيمة المتجمعة الصاعدة لكل خلية ترسم على الرسم وتوصل بخط مستقيم . والحد العلوى الحقيقى للخلية يسمى الإحداثى السينى .

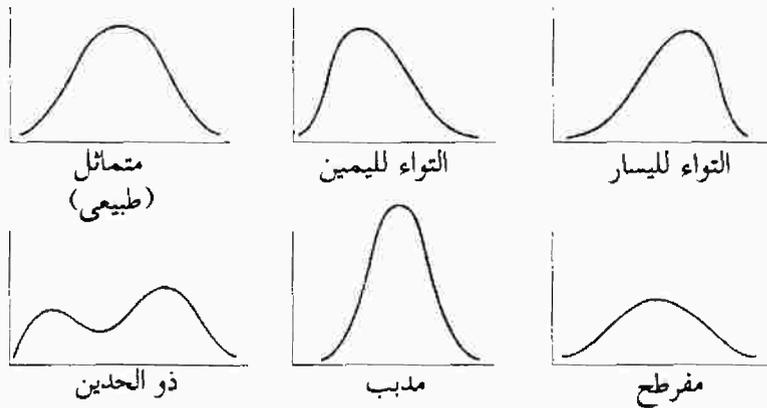
## خواص رسومات التوزيع التكرارى

### Characteristics of Frequency Distribution Graphs

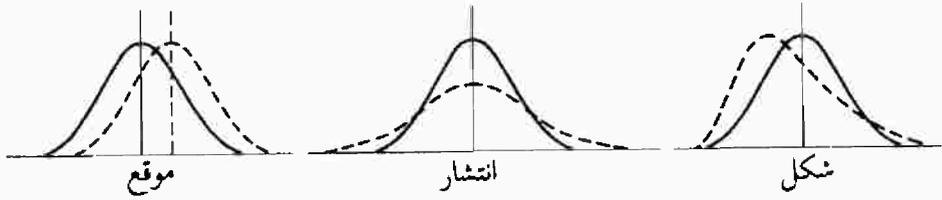
رسومات شكل ٢ - ٦ تستخدم منحنيات ملساء بدلا من المستطيلات المصاحبة للمدرج التكرارى. ويمثل المنحنى الأملس توزيعاً تكرارياً للمجتمع، بينما يمثل المدرج التكرارى التوزيع التكرارى للعينة. والفرق بين المجتمع والعينة يناقش فى قسم آخر فى هذا الفصل.

منحنيات التوزيع التكرارى لها خواص معرفة معينة. إحدى هذه الخواص للتوزيع تهتم التماثل أو عدم التماثل للبيانات. هل البيانات موزعة توزيعاً متساوياً على كلا الجانبين للقيمة المركزية، أم أنها ملتوية ناحية اليمين أو ناحية اليسار؟ خاصية أخرى تهتم عدد المنوالات أو قمم البيانات. ويمكن أن يكون هناك منوالاً واحداً، أو اثنين (توزيع ذات الحدين)، أو عدة منوالات. والخاصية الأخيرة خاصة بمدى حدة البيانات peakedness of the data. فعندما يكون المنحنى حاداً جداً، فيشار إليه بأنه مدبب leptokurtic، وعندما يكون مسطحاً، فيشار إليه بأنه مفرطح platykurtic.

ويمكن أن تعطى التوزيعات التكرارية معلومات كافية عن مشكلة مراقبة الجودة لتوفير أساس لاتخاذ القرار دون أى تحليل إضافى. كما يمكن أن تقارن التوزيعات أيضاً بالنسبة إلى الموقع، والانتشار، والشكل كما هو مبين فى شكل ٢ - ٧.



شكل ٦.٢: خواص التوزيعات التكرارية

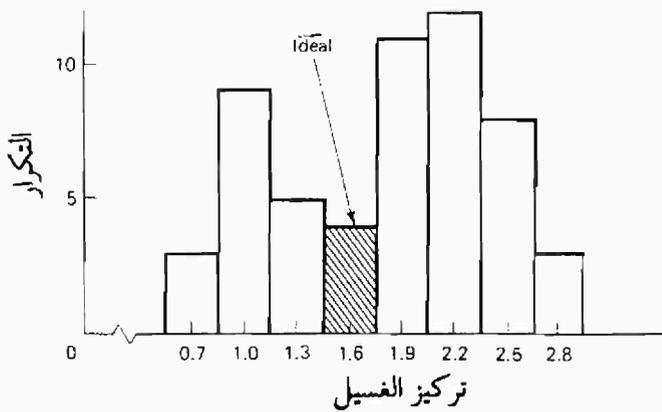


شكل ٧.٢: اختلافات بسبب الموقع، والانتشار، والشكل

### Analysis of Histograms

### تحليل المدرجات التكرارية

يمكن أن يوفر تحليل المدرج التكراري معلومات خاصة بالمواصفات، وشكل التوزيع التكراري للمجتمع، ومشكلة مراقبة جودة محددة. ويبين شكل ٢ - ٨ مدرجا تكراريا للنسبة المئوية لتركيز الغسيل في عملية تنظيف أنبوب من الصلب قبل دهانها. ويقع التركيز المثالي بين 1.45% و 1.74% ، كما هو مبين بالمستطيل المهشور. والتركيز أقل من 1.45% ينتج جودة ضعيفة، أما التركيز أكبر من 1.75% ، بينما ينتج جودة أكثر من اللازم، يكون مكلفا وبالتالي يقلل الإنتاجية. ولايلزم إحصاء معقد لتوضيح أن المقاييس التصحيحية تلزم لجعل انتشار التوزيع أقرب إلى القيمة المثالية 1.6% .



شكل ٨.٢: مدرج تكراري لتركيز الغسيل

## Final Comments

## تعليقات نهائية

هناك نوع آخر للتوزيعات يشبه المدرج التكرارى وهو رسم باريتو Pareto . ولمناقشة هذا النوع من التوزيعات ينصح القارئ بالرجوع إلى الفصلين التاسع والثانى عشر من هذا الكتاب. وتحليل باريتو هو طريقة فعالة جدا لتحديد موقع المشاكل الرئيسية للجودة. والاختلافات بين رسم باريتو والتوزيع التكرارى مزدوجة. فتستخدم الفئات للمحور السينى بدلا من قيم البيانات وتكون الفئات فى ترتيب تنازلى من أعلى تكرر إلى الأقل بدلا من الترتيب العدى.

أحد القيود للتوزيع التكرارى هو أن الحقيقة بأنه لا يبين الترتيب الذى أنتجت به البيانات. وفى كلمات أخرى، يمكن أن توضع البيانات الابتدائية كلها فى جانب واحد والبيانات التى تأتى فيما بعد فى الجانب الآخر. وعندما يحدث هذا الموقف، يصبح تفسير التوزيع التكرارى صعبا. وخريطة الدورة run chart ، التى تناقش فى نهاية الفصل الثالث، تبين الترتيب الذى أنتجت به البيانات ويمكن أن تساعد فى هذا التحليل.

## مقاييس النزعة المركزية MEASURES OF CENTRAL TENDENCY

يكون التوزيع التكرارى كافيا للعديد من مشاكل مراقبة الجودة. إلا أنه مع المدى الواسع للمشاكل تصيح طريقة الرسومات غير مرغوب فيها أو أنها تحتاج إلى معلومات إضافية تقدم بواسطة الطرق التحليلية. وتميز الطرق التحليلية لوصف تجميع بيانات بأنها تشغل مكانا أقل من الرسم. كما أنها تتميز بالسماح بالمقارنات بين تجميع البيانات. كما أنها تسمح أيضا بحسابات واستدلالات إضافية. وهناك طريقتان تحليليتان أساسيتان لوصف تجميع البيانات - قياس النزعة المركزية وقياس التشتت. والقياس الأخير موصوف فى القسم التالى، بينما يغطى القسم الحالى مقاييس النزعة المركزية.

مقياس النزعة المركزية measure of central tendency لتوزيع يكون قيمة عديدة تصف الموقع المركزى للبيانات أو كيف تميل البيانات إلى البناء فى المركز. وهناك ثلاثة مقاييس شائعة الاستخدام: (١) المتوسط، (٢) والوسيط، (٣) والمنوال.

## Average

## المتوسط

المتوسط هو مجموع الملاحظات مقسوما على عددها. وهو المقياس الأكثر استخداما للنزعة المركزية. وهناك ثلاث طرق مختلفة متاحة لحساب المتوسط: (١) البيانات غير المجمعة، (٢) والبيانات المجمعة، (٣) والمتوسط المرجح.

١- البيانات غير المجمعة ungrouped data . تستخدم هذه الطريقة عندما لا تكون البيانات منظمة. ويمثل المتوسط بالرمز  $\bar{X}$ ، ويقرأ "X bar" ويعطى من العلاقة:

$$\bar{X} = \frac{\sum_{i=1}^n X_i}{n} = \frac{X_1 + X_2 + \dots + X_n}{n}$$

حيث:  $\bar{X}$  = المتوسط

$n$  = عدد القيم الملاحظة

$X_1, X_2, \dots, X_n$  = قيمة ملاحظة معرفة بواسطة الدليل

$\sum$  = رمز يعنى «مجموع»

التعبير الأول هو طريقة مبسطة لكتابة الصيغة حيث  $\frac{\sum_{i=1}^n X_i}{n}$  تقرأ «المجموع من 1 إلى n لـ X ذات الدليل i» ويعنى جمع قيم الملاحظات مع بعضها البعض.

## EXAMPLE PROBLEM

## مثال لمشكلة

يختبر أحد القائمين بالفحص قيمة مقاومة خمسة ملفات ويسجل القيم بالآوم ( $\Omega$ )، وكانت القيم كما يلي:

$$X_1 = 3.35, X_2 = 3.37, X_3 = 3.28, X_4 = 3.34, X_5 = 3.30$$

## احسب المتوسط

$$\begin{aligned}\bar{x} &= \frac{\sum_{i=1}^n X_i}{n} \\ &= \frac{3.35 + 3.37 + 3.28 + 3.34 + 3.30}{5} \\ &= 3.33 \Omega\end{aligned}$$

لدى معظم الآلات الحاسبة المقدرة على حساب المتوسط تلقائيا بعد إدخال البيانات.

٢- البيانات المجمعة grouped data . عندما تجمع البيانات في توزيع تكرارى، تسرى الطريقة التالية. صيغة متوسط البيانات المجمعة هي :

$$\bar{x} = \frac{\sum_{i=1}^h f_i X_i}{f_1 + f_2 + \dots + f_h} = \frac{f_1 X_1 + f_2 X_2 + \dots + f_h X_h}{f_1 + f_2 + \dots + f_h}$$

حيث  $n =$  مجموع التكرارات

$f_i =$  تكرار في خلية أو تكرار قيمة ملاحظة

$X_i =$  مركز الخلية أو قيمة ملاحظة

$h =$  عدد الخلايا أو عدد القيم الملاحظة

وتسرى هذه الصيغة عندما يكون التجميع طبقا للخلايا التي بكل خلية أكثر من قيمة ملاحظة واحدة كما هو موضح في مشكلة قضبان الصلب (جدول ٢ - ٦). كما أنها تسرى أيضا عندما يكون لكل قيمة ملاحظة،  $X_i$ ، تكرارها الخاص بها،  $f_i$ ، كما هو مبين في مشكلة أجراس الدخان (جدول ٢ - ١). في هذه الحالة،  $h$  تكون عدد القيم الملاحظة.

وفي كلمات أخرى، إذا ما تم تجميع التوزيع التكرارى فى خلايا، فإن  $X_i$ ، تكون مركز الخلية، وتكون  $f_i$  عدد الملاحظات فى الخلية. وإذا ما تم تجميع التوزيع التكرارى طبقا للقيم الملاحظة الفردية، فتكون  $X_i$  هى القيمة الملاحظة، وتكون  $f_i$  هى عدد مرات حدوث القيمة الملاحظة فى البيانات. وهذه العملية صحيحة لكل من المتغيرات الوثابة والمستمرة.

ويستخدم مركز كل خلية كقيمة ممثلة لهذه الخلية. ويضرب مركز الخلية فى تكرار الخلية، ويجمع حاصل الضرب للخلايا كلها، ويقسم ناتج الجمع على إجمالى عدد الملاحظات. وفى مثال المشكلة المعطاه أدناه، أول ثلاثة أعمدة هى لتوزيع تكرارى تقليدى. والعمود الرابع مستخلص من ضرب العمود الثانى (المركز) والعمود الثالث (التكرار) ويسمى "fx".

### EXAMPLE PROBLEM

### مثال لمشكلة

بمعرفة التوزيع التكرارى لعمر 320 إطار سيارة بالألف كيلومتر (621.37 ميل) كما هو مبين فى جدول ٢ - ٧، فالمطلوب تحديد المتوسط.

$$\bar{x} = \frac{\sum_{i=1}^h f_i X_i}{n}$$

$$= \frac{11,549}{320}$$

$$= 36.1 \quad \text{وهى بالألف كيلومتر}$$

$$\bar{x} = 36.1 \times 10^3 \text{ km}$$

لذلك فإن

جدول ٧.٢: التوزيعات التكرارية لعمر 320 إطار سيارة بالألف كيلو متر

| الحدود    | نقطة المركز<br>$X_i$ | التكرار<br>$f_i$ | حساب قيمه<br>$f_i X_i$  |
|-----------|----------------------|------------------|-------------------------|
| 23.6-26.5 | 25.0                 | 4                | 100                     |
| 26.6-29.5 | 28.0                 | 36               | 1,008                   |
| 29.6-32.5 | 31.0                 | 51               | 1,581                   |
| 32.6-35.5 | 34.0                 | 63               | 2,142                   |
| 35.6-38.5 | 37.0                 | 58               | 2,146                   |
| 38.6-41.5 | 40.0                 | 52               | 2,080                   |
| 41.6-44.5 | 43.0                 | 34               | 1,462                   |
| 44.6-47.5 | 46.0                 | 16               | 736                     |
| 47.6-50.5 | 49.0                 | 6                | 294                     |
| الاجمالي  |                      | $n = 320$        | $\sum f_i X_i = 11,549$ |

عند مقارنة متوسط محسوب بهذه الطريقة مع متوسط محسوب باستخدام طريقة البيانات غير المجمعة، يمكن أن يوجد اختلاف بسيط. وحدث هذا الاختلاف من أن الملاحظات في كل خلية تكون موزعة توزيعاً غير منتظماً في الخلية. وفي العملية الحقيقية لا يكون للاختلاف قيمة كافية للتأثير على دقة المشكلة.

٣- المتوسط المرجح  $\text{weighted average}$ . عندما يدمج عدد من المتوسطات لها تكرارات مختلفة، فيحسب متوسط مرجح  $\text{weighted average}$ . وصيغة حساب المتوسط المرجح هي كما يلي :

$$\bar{x}_w = \frac{\sum_{i=1}^n w_i \bar{x}_i}{\sum_{i=1}^n w_i}$$

حيث:  $\bar{x}_w$  = المتوسط المرجح

$w_i$  = ترجيح المتوسط  $i$

## EXAMPLE PROBLEM

## مثال لمشكلة

تؤدي اختبارات الشد لسبيكة ألومونيوم في ثلاثة أوقات مختلفة، ينتج عنها ثلاث قيم متوسطات مختلفة بالميجا بسكال (MPa). في الحالة الأولى تؤدي خمسة اختبارات بمتوسط 207 ميجا بسكال، وفي الحالة الثانية تؤدي ستة اختبارات بمتوسط 203 ميجا بسكال، وفي الحالة الأخيرة تؤدي ثلاثة اختبارات بمتوسط 206 ميجا بسكال. حدد المتوسط المرجح.

$$\begin{aligned}\bar{x}_w &= \frac{\sum_{i=1}^n w_i \bar{x}_i}{\sum_{i=1}^n w_i} \\ &= \frac{(5)(207) + (6)(203) + (3)(206)}{5 + 6 + 3} \\ &= 205 \text{ MPa}\end{aligned}$$

طريقة المتوسط المرجح هي حالة خاصة لطريقة البيانات المجمعة حيث لا تكون البيانات منظمة في توزيع تكرارى. وفي المثال السابق كانت الترجيحات أرقاماً صحيحة. وطريقة أخرى لحل نفس المشكلة تكون باستخدام النسب. لهذا،

$$w_1 = \frac{5}{5 + 6 + 3} = 0.36$$

$$w_2 = \frac{6}{5 + 6 + 3} = 0.43$$

$$w_3 = \frac{3}{5 + 6 + 3} = 0.21$$

ويكون مجموع الترجيحات مساوياً 1.00. الطريقة الأخيرة يمكن أن تكون ضرورية عندما تعطى الترجيحات كنسب مئوية أو ككسور عشرية.

ودون ذكر أى شئ آخر، فتعنى  $\bar{x}$  متوسط القيم الملاحظة  $\bar{x}_2$ . وتستخدم نفس المعادلة في ايجاد :

$\bar{x}$  أو  $X_{\bar{x}}$  - متوسط المتوسطات

$\bar{r}$  - متوسط المدى

$\bar{c}$  - متوسط العدد غير المطابق

$\bar{s}$  - متوسط الانحراف المعياري للعينة، الخ

## Median

## الوسيط

مقياس آخر للترعة المركزية هو الوسيط median ، والذي يعرف كقيمة تقسم سلسلة من الملاحظات المرتبة بحيث أن عدد العناصر فوقها يساوى عدد العناصر تحتها.

١- طريقة البيانات غير المجمعة ungrouped technique . هناك حالتان ممكنتان في تحديد الوسيط لسلسلة من البيانات غير المجمعة - عندما يكون العدد في السلسلة فرديا، يكون الوسيط مركز هذه القيم. لهذا، فئة الأعداد المرتبة 3 و 4 و 5 و 6 و 8 و 8 و 10 لها الوسيط 6 وفئة الأعداد المرتبة 22 و 24 و 24 و 24 و 30 يكون لها الوسيط. وعندما يكون العدد في السلسلة زوجيا، يكون الوسيط متوسط العددين الموجودين في المنتصف. لهذا، فئة الأعداد المرتبة 3 و 4 و 5 و 6 و 8 و 8 و 8 يكون لها الوسيط عبارة عن متوسط 5 و 6 أى 5.5 . فإذا كان العدداً المتوسطان مثل بعضهما كما في فئة الأعداد المرتبة 22 و 24 و 24 و 24 و 30 فيكون الوسيط مساويا لأى منهما أى مساويا 24 فى هذه الحالة. ويجب أن يحتاط القارئ من أن الأعداد يجب أن تكون مرتبة قبل حساب الوسيط.

٢- طريقة البيانات المجمعة grouped technique . عندما تكون البيانات مجمعة فى توزيع تكرارى، فيتم الحصول على الوسيط عن طريق إيجاد الخلية التى لها رقم متوسط ثم يستكمل من داخل الخلية. وصيغة الاستكمال هى كما يلي :

$$Md = L_m + \left( \frac{\frac{n}{2} - cf_m}{f_m} \right) i$$

حيث :  $Md$  = الوسيط

$L_m$  = الحد السفلى للخلية مع الوسيط

$n$  = إجمالي عدد الملاحظات

$cf_m$  = التكرار المتجمع الصاعد لكل الخلايا الأقل من  $L_m$

$f_m$  = تكرار خلية الوسيط

$i$  = فترة الخلية

لتوضيح استخدام هذه الصيغة، تستخدم البيانات الموجودة في جدول ٢ - ٧. بالعد من أقل خلية (مركزها 25.0)، يتم الوصول إلى نقطة منتصف الطريق (320/2) (160) = في الخلية بقيمة مركز 37.0 وأدنى حد 35.6. التكرار المتجمع الصاعد ( $cf_m$ ) يكون 154، وتكون فترة الخلية 3، وتكرار خلية الوسيط يكون 58

$$\begin{aligned} Md &= L_m + \left( \frac{\frac{n}{2} - cf_m}{f_m} \right) i \\ &= 35.6 + \left( \frac{\frac{320}{2} - 154}{58} \right) 3 \\ &= 35.9 \quad (\text{which is in } 1000 \text{ km}) \end{aligned}$$

إذا بدأ العد في قمة التوزيع، فيعد التكرار المتجمع الصاعد لخلية الحد العلوى وتطرح الكمية المستكملة من الحد العلوى. إلا أنه من الأكثر شيوعا البدء في العد في قاع التوزيع.

## Mode

## المنوال

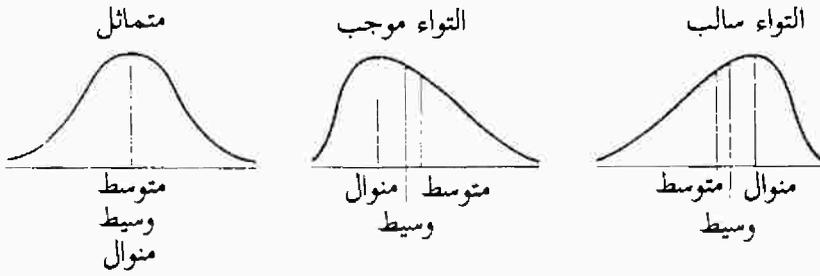
المنوال (MO) mode لفئة أعداد هو القيمة التي تحدث بأكثر تكرار. ومن الممكن للمنوال ألا يوجد في سلسلة من الأعداد أو أنه يكون له أكثر من قيمة واحدة. للتوضيح سلسلة الأعداد 3 و 3 و 4 و 5 و 5 و 5 و 7 وسلسلة الأعداد 22 و 23 و 25 و 30 و 32 و 36 ليس لها منوال وسلسلة الأعداد 104 و 105 و 105 و 107 و 108 و 109 و 109 و 110 و 211 لها منوالان، هما 105 و 109. ويشار إلى سلسلة الأعداد بأنها أحادية المنوال unimodal إذا كان لها منوال واحد فقط، وبأنها مزدوجة المنوال bimodal إذا كان لها منوالان، ومتعددة المنوال multimodal إذا كان لها أكثر من منوالين.

وعندما تجمع البيانات في توزيع تكرارى، فإن مركز الخلية التي لها أعلى تكرار يكون هو المنوال، حيث أن هذه النقطة تمثل أعلى نقطة (أكبر تكرار) للمدرج التكرارى. ومن الممكن الحصول على تقدير أفضل للمنوال بالاستكمال بطريقة تشبه الطريقة المستخدمة مع الوسيط، إلا أن هذا لا يكون ضروريا، حيث أن المنوال يستخدم أساسا كطريقة فحص لتحديد النزعة المركزية، ولا تلزم دقة أكبر من مركز الخلية.

## العلاقة بين مقاييس النزعة المركزية

## Relationship Among the Measures of Central Tendency

الاختلافات بين الثلاثة مقاييس للنزعة المركزية مبينة في المصطلح التكرارى الأملس الموجود في شكل ٢ - ٩. عندما يكون التوزيع متماثلا، تكون قيم المتوسط، والوسيط، والمنوال متطابقة، وعندما يكون التوزيع منحرفا، تختلف القيم.



شكل ٩.٢: العلاقة بين المتوسط والوسيط، والمنوال

المتوسط هو المقياس الأكثر استخداماً للنزعة المركزية. ويستخدم عندما يكون التوزيع متماثلاً أو غير مميز انحرافه لليمين أو اليسار، وعندما يراد حساب الإحصاء الإضافي مثل قياسات التشتت، وخرائط المراقبة، وغيرها اعتماداً على المتوسط، وعندما تلزم قيمة مستقرة للإحصاء الاستدلالي.

ويصبح الوسيط مقياساً فعالاً للنزعة المركزية عندما يكون التوزيع ملتويًا موجبا (لليمين) أو سالبا (لليسار). ويستخدم عندما يكون مطلوب قيمة مركز التوزيع بالضبط. وعندما يكون للتوزيع قيم قصوى، فيتأثر المتوسط بطريقة معكوسة بينما يظل الوسيط دون تغيير. لهذا، في سلسلة أعداد مثل 12 و 13 و 14 و 15 و 16 يكون الوسيط والمتوسط متطابقين ومتساويين 14. إلا أنه إذا تغيرت القيمة الأولى إلى 2، فيظل الوسيط كما هو عند 14، بينما يصبح المتوسط 12.

ويستخدم المنوال عندما يراد مقياس سريع وتقريبي للنزعة المركزية. لهذا، يوجد منوال المدرج التكراري بسهولة عن طريق الفحص البصري. بالإضافة إلى ذلك، يستخدم المنوال في وصف القيم التقليدية أكثر للتوزيع، مثل منوال عمر مجموعة معينة.

وهناك مقاييس أخرى للنزعة المركزية وهي الوسط الهندسي، والوسط التوافقي، والوسط التريبي. ولا تستخدم هذه المقاييس في مراقبة الجودة.

## MEASURES OF DISPERSION

## مقاييس التشتت

### Introduction

### مقدمة

لقد نوقشت في القسم السابق طرق وصف النزعة المركزية للبيانات. ووسيلة أخرى للإحصاء تتكون من مقاييس التشتت measures of dispersion، والتي تصف كيفية انتشار البيانات للخارج أو بعثرتها على كل جانب من جوانب القيمة المركزية. وتلزم قياسات التشتت وقياسات النزعة المركزية لوصف تجميع من البيانات. لتوضيح ذلك، العاملون في أقسام الطلاء المعدني والتجميع لأحد المصانع يكون لهم متوسط أجر أسبوعي متطابق قيمته 225.36 دولار، إلا أن قسم الطلاء المعدني له قيمة عليا مقدارها 230.27 دولار وقيمة دنيا مقدارها 219.43 دولار، بينما قسم التجميع له قيمة عليا مقدارها 280.79 دولار، وقيمة دنيا مقدارها 173.54 دولار. بيانات قسم التجميع تكون منتشرة للخارج أو مشتتة أكثر من المتوسط عن بيانات قسم الطلاء المعدني.

ومقاييس التشتت المناقشة في هذا القسم هي المدى، والانحراف المعياري، والتباين. والمقاييس الأخرى مثل وسط التباين الحسابي والانحراف الربيعي لا تستخدم في مراقبة الجودة.

### Range

### المدى

المدى range لسلسلة من الأعداد هو الفرق بين أكبر وأقل قيمة أو ملاحظة. ورمزيا، فإنه يتحدد بالعلاقة:

$$R = X_H - X_L$$

حيث:  $R =$  المدى

$X_H =$  أعلى ملاحظة في السلسلة

$X_L =$  أقل ملاحظة في السلسلة

### EXAMPLE PROBLEM

### مثال لمشكلة

إذا كان أعلى أجر أسبوعي في قسم التجميع هو 280.79 دولار وأقل أجر أسبوعي هو 173.54 دولار، حدد المدى.

$$\begin{aligned} R &= X_H - X_L \\ &= \$280.79 - \$173.54 \\ &= \$107.25 \end{aligned}$$

والمدى هو الأبسط والأسهل في حسابه بالنسبة إلى مقياس التشتت. وهناك مقياس مرتبط به ويستخدم في بعض الأحيان وهو مركز المدى midrange ، وهو عبارة عن خارج قسمة المدى على 2 .

### Standard Diviation

### الانحراف المعياري

الانحراف المعياري standard diviation هو قيمة عددية بوحدات القيم الملاحظة والتي تقيس انتشار نزعة البيانات. ويبين الانحراف المعياري الكبير مقدرة أكبر على التغير للبيانات عن الانحراف المعياري الصغير. وبالرموز فإنه يعطى بالصيغة التالية :

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

حيث :  $S =$  الانحراف المعياري للعينة

$X_i =$  القيمة الملاحظة

$\bar{X} =$  المتوسط

$n =$  عدد القيم الملاحظة

يستخدم جدول ٢ - ٨ في توضيح مفهوم الانحراف المعياري. العمود الأول ( $X_i$ ) يعطى ست قيم ملاحظة بالكيلوجرام، ومن هذه القيم يتم الحصول على المتوسط  $\bar{X} = 3.0$  العمود الثاني ( $X_i - \bar{X}$ ) ، هو انحراف القيم الملاحظة الفردية عن المتوسط. إذا ما جمعنا الانحرافات، تكون النتيجة صفراً، وهذا هو الحال دائماً، إلا أن هذا لا يقود إلى قياس التشتت. إلا أنه إذا تم تربيع الانحرافات، فإنها تصبح موجبة ويصبح مجموعها أكبر من صفر. والحسابات مبينة في العمود الثالث،  $(X_i - \bar{X})^2$  ، مع المجموع الناتج مساوياً 0.08 ، والذي يتغير طبقاً للقيم الملاحظة. ومتوسط مربعات الانحرافات يمكن أن يوجد عن طريق القسمة على  $n$  ، إلا أنه لأسباب نظرية فإننا نقسم على  $n-1$ <sup>(١)</sup>. لهذا،

(١) سبب استخدام  $n-1$  هو لأن درجة حرية واحدة تفقد بسبب استخدام عينة إحصائية  $\bar{X}$  ، بدلا من استخدام معلمة المجتمع،  $\mu$ .

$$\frac{\sum (X_i - \bar{x})^2}{n - 1} = \frac{0.08}{6 - 1} = 0.016 \text{ kg}^2$$

والتي تعطي نتيجة لها وحدات مربعة. هذه النتيجة ليست مقبولة كمقياس للتشتت، إلا أنها مفيدة كمقياس للقابلية للتغير للإحصاء المتقدم.

جدول ٨.٢: تحليل الانحراف المعياري

| $X_i$           | $X_i - \bar{X}$ | $(X_i - \bar{X})^2$ |
|-----------------|-----------------|---------------------|
| 3.2             | +0.2            | 0.04                |
| 2.9             | -0.1            | 0.01                |
| 3.0             | 0.0             | 0.00                |
| 2.9             | -0.1            | 0.01                |
| 3.1             | +0.1            | 0.01                |
| 2.9             | -0.1            | 0.01                |
| $\bar{X} = 3.0$ | $\Sigma = 0$    | $\Sigma = 0.08$     |

ويسمى تباين variance ويعطى له الرمز  $s^2$ . فإذا ما أخذنا الجذر التربيعي، فتصبح النتيجة بنفس وحدات القيم الملاحظة. وتكون الحسابات كما يلي :

$$S = \sqrt{\frac{\sum (X_i - \bar{x})^2}{n - 1}} = \sqrt{\frac{0.08}{6 - 1}} = 0.13 \text{ kg}$$

هذه الصيغة للتوضيح بدلا من كونها لغرض إجراء الحسابات. ونظرا لأن صيغة البيانات يمكن أن تكون مجمعة أو غير مجمعة، فتوجد طرق حسابية مختلفة.

١ - طريقة البيانات غير المجمعة ungrouped technique . الصيغة المستخدمة في تعريف الانحراف المعياري يمكن أن تستخدم للبيانات غير المجمعة. إلا أنه هناك صيغة أكثر راحة تستخدم في أغراض الحسابات :

$$S = \sqrt{\frac{n \sum_{i=1}^n X_i^2 - \left( \sum_{i=1}^n X_i \right)^2}{n(n-1)}}$$

### EXAMPLE PROBLEM

### مثال لمشكلة

حدد الانحراف المعياري لمحتوى الرطوبة لقضيب من الورق المقوى. نتائج ست قراءات لرقائق الورق كانت كما يلي : 6.7% و 6.0% و 6.4% و 6.4% و 5.9% و 5.8% .

$$\begin{aligned} S &= \sqrt{\frac{n \sum_{i=1}^n X_i^2 - \left( \sum_{i=1}^n X_i \right)^2}{n(n-1)}} \\ &= \sqrt{\frac{6(231.26) - (37.2)^2}{6(6-1)}} \\ &= 0.35\% \end{aligned}$$

بعد إدخال البيانات، تحسب العديد من الآلات الحاسبة الانحراف المعياري طبقا للطلب.

٢ - طريقة البيانات المجمعة grouped data . عندما تكون البيانات مجمعة في توزيع تكرارى، تطبق الطريقة التالية. صيغة الانحراف المعياري للبيانات المجمعة هي كما يلي :

$$S = \sqrt{\frac{n \sum_{i=1}^h (f_i X_i^2) - \left( \sum_{i=1}^h f_i X_i \right)^2}{n(n-1)}}$$

حيث الرموز  $f_i$  و  $x_i$  و  $n$  و  $h$  يكون لها نفس المعنى الذى سبق إعطاؤه لمتوسط البيانات المجمعة.

ولاستخدام هذه الطريقة، يضاف عمودان إضافيان للتوزيع التكرارى. ويسمى هذان العمودان "fX" و "fX<sup>2</sup>" ، كما هو مبين فى جدول ٢ - ٩. ويجب تذكر أن عمود "fX" يلزم لحسابات المتوسط، لهذا فإنه يلزم عمود إضافي وحيد لحساب الانحراف المعياري. والطريقة مبينة بمثال المشكلة التالى :

**جدول ٩.٢:** سرعات سيارات الركاب (كم/ساعة) أثناء فترة من 15 دقيقة.

| الحدود      | نقطة المركز | التكرار  | الحسابات           |                         |
|-------------|-------------|----------|--------------------|-------------------------|
|             | $X_i$       | $f_i$    | $f_i X_i$          | $f_i X_i^2$             |
| 72.6-81.5   | 77.0        | 5        | 385                | 29,645                  |
| 81.6-90.5   | 86.0        | 19       | 1634               | 140,524                 |
| 90.6-99.5   | 95.0        | 31       | 2945               | 279,775                 |
| 99.6-108.5  | 104.0       | 27       | 2808               | 292,032                 |
| 108.6-117.5 | 113.0       | 14       | 1582               | 178,766                 |
| Total       |             | $n = 96$ | $\Sigma fX = 9354$ | $\Sigma fX^2 = 920,742$ |

### EXAMPLE PROBLEM

### مثال لمشكلة

بمعرفة التوزيع التكرارى الموجود فى جدول ٩.٢ لسرعات سيارات الركاب أثناء فترة من 15 دقيقة، حدد المتوسط والانحراف المعياري.

$$\bar{x} = \frac{\sum_{i=1}^h f_i X_i}{n}$$

$$S = \sqrt{\frac{n \sum_{i=1}^h (f_i X_i^2) - \left( \sum_{i=1}^h f_i X_i \right)^2}{n(n-1)}}$$

$$= \frac{9354}{96}$$

$$= 97.4 \text{ km/h}$$

$$= \sqrt{\frac{96(920,742) - (9354)^2}{96(96-1)}}$$

$$= 9.9 \text{ km/h}$$

لا تقرب  $\sum fx$  أو  $\sum fx^2$  حيث أن هذا الإجراء يؤثر على الدقة. ولدى معظم الآلات الحاسبة المقدرة على إدخال بيانات مجمعة وحساب  $s$  طبقا للطلب.

وبدون أن يذكر أى شيء خلاف ذلك، فإن  $s$  تعنى  $s_x$ ، أى الانحراف المعياري لعينة القيم الملاحظة. وتستخدم نفس الصيغة في إيجاد :

$S_x$  - الانحراف المعياري لمتوسطات العينات

$S_p$  - الانحراف المعياري لنسب العينات

$S_R$  - الانحراف المعياري لمجموعات مدى العينات

$S_S$  - الانحراف المعياري للانحرافات المعيارية للعينات، إلخ.

والانحراف المعياري هو قيمة دليل يقيس التشتت في البيانات. ومن الأفضل رؤيته على أنه فهرس يكون معرفا بواسطة الصيغة. وكلما قلت قيمة الانحراف المعياري كلما تحسنت الجودة، حيث أن التوزيع يكون أكثر انضغاطا حول القيمة المركزية. وكذلك، يساعد الانحراف المعياري في تعريف المجتمعات.

### العلاقة بين مقاييس التشتت

#### Relationship Between the Measures of Dispersion

في مراقبة الجودة يكون المدى مقياسا شائعا جدا للتشتت، ويستخدم في أحد خرائط المراقبة الرئيسية. وتكمن الميزة الأولية للمدى في تقديمه معرفة بإجمالي انتشار البيانات. كما أنه مفيد أيضا عندما تكون كمية البيانات صغيرة جدا أو مبعثرة جدا لتبرير حساب مقياس أكثر دقة للتشتت. والمدى ليس وظيفة قياس للنزعة المركزية. فمع تزايد عدد الملاحظات، تقل دقة المدى، حيث أنه يصبح من الأسهل حدوث قراءات كبيرة جدا وأخرى صغيرة جدا. ويقترح أن يكون استخدام المدى محددًا بعشر ملاحظات كحد أقصى.

ويستخدم الانحراف المعياري عندما يكون مطلوباً مقياساً أكثر إتقاناً. كما أنه المقياس الأكثر شيوعاً للتشتت في عمل مراقبة الجودة ويستخدم عندما يراد عمل حسابات إحصائية تالية. عندما يكون للبيانات قيمة مرتفعة ارتفاعاً غير عادي أو منخفضة انخفاضاً غير عادي، فيكون الانحراف المعياري مرغوباً أكثر عن المدى.

## OTHER MEASURES

## مقاييس أخرى

هناك مقياسان آخران يتكرر استخدامهما في تحليل مجموعة من البيانات - الالتواء والتفرطح.

### Skewness

### الالتواء

كما سبق ذكره، الالتواء skewness هو فقدان التماثلية للبيانات. والصيغة معطاه كما يلي<sup>(١)</sup>:

$$\alpha_3 = \frac{\sum_{i=1}^h f_i (X_i - \bar{x})^3 / n}{S^3}$$

حيث  $\alpha_3$  تمثل الالتواء.

والالتواء عبارة عن عدد يذكر حجمه مدى الترحيل من التماثلية. إذا كانت قيمة  $\alpha_3$  هي صفر، تكون البيانات متماثلة، وإذا كانت أكبر من صفر (موجبة)، تكون البيانات ملتوية لليمين، وهذا يعني وجود ذيل طويل ناحية اليمين، وإذا كانت أقل من صفر (سالبة)، فتكون البيانات ملتوية ناحية اليسار، وهذا يعني وجود ذيل طويل ناحية اليسار. انظر شكل ٢ - ٩ لرؤية تمثيل بياني للالتواء. وتشمل قيم +1 أو -1 توزيعاً متماثلاً تماماً تقريباً.

(١) هذه الصيغة عبارة عن تقريب جيد بدرجة كافية لاستخدامها في معظم الأغراض.

## EXAMPLE PROBLEM

## مثال لمشكلة

حدد التواء التوزيع التكرارى لجدول ٢ - ١٠. حسب المتوسط وكذلك الانحراف المعياري وكانت قيمتهما 7.0 و 2.32 ، على التوالي .

$$\alpha_3 = \frac{\sum_{i=1}^h f_i (X_i - \bar{X})^3 / n}{S^3}$$

$$= \frac{-162 / 31}{2.32^3}$$

$$= -0.42$$

جدول ١٠.٢: بيانات لمثال مشكلة الالتواء والتفرطح

| $X_i$         | $f_i$ | $X_i - \bar{X}$ | $f_i(X_i - \bar{X})^3$ | $f_i(X_i - \bar{X})^4$ |
|---------------|-------|-----------------|------------------------|------------------------|
| 1             | 1     | $(1 - 7) = -6$  | $1(-6)^3 = -216$       | $1(-6)^4 = 1296$       |
| 4             | 6     | $(4 - 7) = -3$  | $6(-3)^3 = -162$       | $6(-3)^4 = 486$        |
| 7             | 16    | $(7 - 7) = 0$   | $16(0)^3 = 0$          | $16(0)^4 = 0$          |
| 10            | 8     | $(10 - 7) = +3$ | $8(+3)^3 = +216$       | $8(+3)^4 = 648$        |
| $\Sigma = 31$ |       |                 | $\Sigma = -162$        | $\Sigma = 2430$        |

قيمة الالتواء 0.42 - تذكر لنا أن البيانات ملتوية جدا إلى اليسار. ويمكن أن يحدد الفحص البصري لعمودى  $f$  و  $x$  أو للمدرج التكرارى نفس المعلومات.

ولكى تستخدم قيمة الالتواء، يجب أن تكون قيمة  $n$  كبيرة، ولتكن 100 على الأقل. كذلك، يجب أن يكون التوزيع أحادى النوال. وتقدم قيمة الانحراف معلومات خاصة بشكل توزيع المجتمع. مثال ذلك، يكون للتوزيع الطبيعي التواء صفر.

## Kurtosis

## التفرطح

كما سبق ذكره من قبل التفرطح هو حالة قمة البيانات. وتعطى الصيغة على النحو التالي<sup>(١)</sup>:

$$\alpha_4 = \frac{\sum_{i=1}^h f_i (X_i - \bar{X})^4 / n}{S^4}$$

حيث  $\alpha_4$  تمثل التفرطح

التفرطح هو قيمة بدون وحدات وتستخدم كمقياس مقارن لارتفاع القمة في توزيعين. انظر شكل ٢ - ٦ وشكل ٢ - ١١ لترى التمثيل البياني للتفرطح.

## EXAMPLE PROBLEM

## مثال لمشكلة

حدد التفرطح للتوزيع التكرارى لجدول ٢ - ١٠، والذي له  $\bar{X} = 7.0$  و  $S =$

2.32.

$$\begin{aligned} \alpha_4 &= \frac{\sum_{i=1}^h f_i (X_i - \bar{X})^4 / n}{S^4} \\ &= \frac{2430 / 31}{2.32^4} \\ &= 2.70 \end{aligned}$$

قيمة التفرطح لا توفر أى معلومات فى حد ذاتها - فيجب أن تقارن مع توزيع آخر. واستخدام قيمة التفرطح هو نفسه مثل الانحراف - حجم عينة  $n$  كبير، وتوزيع أحادى المنوال. مثال ذلك، التوزيع الطبيعي، توزيع متوسط التفرطح mesokurtic، له (١) هذه الصيغة هي تقريب جيد بدرجة كافية لتناسب معظم الأغراض.

قيمة تفرطح تساوى 3 ، أى  $a = 3$  . إذا كانت  $a > 3$  فيكون ارتفاع التوزيع له قيمة اكبر من التوزيع الطبيعي ، توزيع مدبب leptokurtic ، وإذا كانت  $a < 3$  فيكون ارتفاع التوزيع أقل فى قمته عن التوزيع الطبيعي ، توزيع مفرطح platykurtic .

مفاهيم الالتواء والتفرطح تكون مفيدة فى أنها توفر بعض المعلومات عن شكل التوزيع .

### مفهوم المجتمع والعينة CONCEPT OF A POPULATION AND A SAMPLE

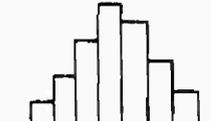
عند هذه النقطة ، يكون من المرغوب فيه فحص مفهوم المجتمع والعينة . ولكى نعد توزيعا تكراريا لأوزان قضبان من الصلب ، يتم اختيار جزء صغير ، أو عينة sample ، لتمثيل كل قضبان الصلب . وبالمثل ، البيانات المجموعة الخاصة بسرعات سيارات الركاب تمثل جزءا صغيرا فقط لكل سيارات الركاب . والمجتمع هو التجميع الكلى للقياسات ، وفى المثال السابق ، يكون المجتمع هو كل قضبان الصلب وكل سيارات الركاب . وعند حساب المتوسطات ، والانحرافات المعيارية ، والمقاييس الأخرى من عينات ، فيشار إليها بأنها إحصائيات statistics . وحيث أن تكوين العينة يتغير ، فتكون الإحصائيات المحسوبة أكبر من أو أقل من قيمها الحقيقية للمجتمع ، أو عن المعلمات (المؤشرات) parameters . وتعتبر المعلمات أنها قيم أدلة ثابتة (نمطية) أو أنها أفضل تقدير لهذه القيم المتاح فى وقت محدد .

ويمكن أن يكون للمجتمع عدد محدد من العناصر ، مثل إنتاج قضبان الصلب اليومى . كما يمكن أن يكون لانهاثيا أو لانهاثيا تقريبا ، مثل عدد مسامير البرشام فى إنتاج الطائرات النفاثة السنوى . ويمكن أن يعرف المجتمع بطريقة مختلفة طبقا لموقف معين . لهذا ، يمكن أن تشمل دراسة المنتج مجتمع إنتاج ساعة واحدة ، أو إنتاج أسبوع ، أو 5000 قطعة ، وغيرها .

وحيث انه من النادر أن يكون ممكنا قياس المجتمع كله، فيتم اختيار عينة، وتكون المعاينة لازمة عندما يكون من المستحيل قياس المجتمع كله، أو عندما تكون تكاليف ملاحظة كل البيانات مرتفعة جدا، أو عندما يدمر الفحص اللازم المنتج، أو عندما يكون اختبار كل المجتمع خطيرا جدا، كما في حالة الأدوية الجديدة. وفعليا، فمحلل المجتمع كله لن يكون أكثر دقة من المعاينة. فقد تبين أن فحص يدوى 100% ليس بنفس دقة المعاينة. وربما يرجع ذلك إلى الحقيقة بأن الضجر والتعب يتسببان في حكم مسبق للعنصر المفحوص بأنه مقبول من قبيل الفاحص.

وعند تحديد مجتمع، يستخدم الحرف اليوناني المناظر. لهذا، يكون لمتوسط العينة الرمز  $\bar{x}$ ، ويكون للمتوسط الحسابي للمجتمع الرمز  $\mu$  (ميو). لاحظ أن كلمة «متوسط» تغيرت إلى «وسط حسابي» عند استخدام المجتمع. والرمز  $\bar{x}_0$  هو القيمة النمطية أو الدليل النمطي. وتبنى المفاهيم الرياضية على  $\mu$ ، وهي القيمة الحقيقية  $\bar{x}_0$  تمثل مكافئ عملي بفرض استخدام المفاهيم. والانحراف المعياري للعينة له الرمز  $s$ ، وللانحراف المعياري للمجتمع الرمز  $\sigma$  (سيجما). والرمز  $s$  هو قيمة نمطية أو إشارة نمطية وله نفس العلاقة مع  $\sigma$  مثل علاقة  $\bar{x}_0$  مع  $\mu$ . وقد لا تعرف قيمة المجتمع الحقيقية على الإطلاق، لهذا، أحيانا يستخدم الرمز  $\hat{\mu}$  و  $\hat{\sigma}$  للدلالة على «تقدير لـ». وتوجد مقارنة بين العينة والمجتمع في جدول ٢ - ١١. وسوف تذكر مقارنات إضافية عند حدوثها فيما بعد.

جدول ١١.٢: مقارنة بين العينة والمجتمع

| المجتمع   | العينة  |
|---|---|
| معلمة   | إحصائية   |
| $\bar{x}_0$ - المتوسط الحسابي   | $\bar{x}$ - المتوسط   |
| $\sigma$ (s <sub>0</sub> ) - الانحراف المعياري                                      | $s$ - الانحراف المعياري للعينة  |
|  |  |

الهدف الأساسى من اختيار العينة هو تعلم شىء ما عن المجتمع والذى يساعد فى اتخاذ نوع ما من القرارات. ويجب أن يكون للعينة المختارة طبيعة أنها تميل إلى تمثيل أو تشابه المجتمع. كيف تمثل العينة المجتمع بنجاح يكون دالة فى حجم العينة، والفرصة، وطريقة المعاينة، وما إذا كانت الشروط تتغير أم لا.

وبين جدول ٢ - ١٢ نتائج تجربة توضح العلاقة بين العينات والمجتمع. ففي الحاوى الذى به 800 كرة زرقاء، و 200 كرة خضراء قطر كل منها 5 مم (حوالى 3/16 بوصة)، تعتبر الكرات البالغ عددها 1000 مجتمعا، به 20% كرات خضراء. وتختار عينات حجم كل منها 10 كرات، لعرضها على المنضدة قبل إعادتها إلى الحاوى. وتوضح المنضدة الاختلافات بين نتائج العينات وما يجب توقعه من المجتمع المعروف. وفى العينتين 2 و 7 فقط كانت إحصائية العينة مساوية لمعلمة المجتمع. ويوجد عامل فرصة أكيد يحدد تكوين العينة. فعندما تدمج العينات الثمان الفردية فى عينة واحدة كبيرة، تصبح نسبة الكرات الخضراء 18.8%، وهى قريبة من قيمة المجتمع 20%.

**جدول ١٢.٢: نتائج ثمان عينات لكرات زرقاء وخضراء تسحب من مجتمع معروف**

| رقم العينة | حجم العينة | عدد الكرات الزرقاء | عدد الكرات الخضراء | النسبة المئوية للكرات الخضراء |
|------------|------------|--------------------|--------------------|-------------------------------|
| 1          | 10         | 9                  | 1                  | 10                            |
| 2          | 10         | 8                  | 2                  | 20                            |
| 3          | 10         | 5                  | 5                  | 50                            |
| 4          | 10         | 9                  | 1                  | 10                            |
| 5          | 10         | 7                  | 3                  | 30                            |
| 6          | 10         | 10                 | 0                  | 0                             |
| 7          | 10         | 8                  | 2                  | 20                            |
| 8          | 10         | 9                  | 1                  | 10                            |
| الاجمالى   | 80         | 65                 | 15                 | 18.8                          |

بينما حدثت استدالات عن المجتمع من العينات، فإنه من الحقيقى أن معرفة المجتمع تقدم معلومات لتحليل العينة. لهذا، يكون ممكنا تحديد ما إذا كانت العينة قد سحبت من مجتمع معين. هذا المفهوم يكون ضروريا فى فهم نظرية خرائط المراقبة. وقد تأجلت مناقشة تفصيلية أكثر إلى الفصل الثالث.

## THE NORMAL CURVE

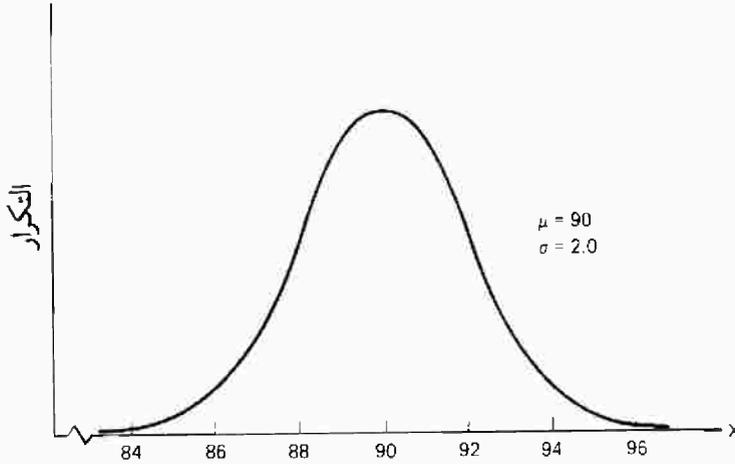
## المنحنى الطبيعي

## Description

## الوصف

بالرغم من وجود مجتمعات مختلفة عديدة مثل وجود الشروط، إلا أنها يمكن أن توصف بواسطة أنواع عامة محددة. أحد أنواع المجتمعات شائعة الاستخدام يسمى بالمنحنى الطبيعي normal curve أو توزيع جاوس Gaussian distribution . والمنحنى الطبيعي هو توزيع متماثل، يشبه الجرس، له منوال واحد، وله نفس القيمة لكل من الوسط الحسابي، والوسيط، والمنوال.

المنحنى الطبيعي أو التوزيع الطبيعي يعد من المدرج التكرارى. وكلما كبر حجم عينة المدرج التكرارى وازداد كبيرا، تصبح فترة الخلية أقل وأكثر قلة. وعندما يكون حجم العينة كبيرا جدا وتصبح فترة الخلية قليلة جدا، يبدو المدرج التكرارى كمضلع تكرارى أملس أم منحنى يمثل المجتمع. المنحنى لمجتمع طبيعى لعدد 1000 ملاحظة للمقاومة الكهربائية لأحد المعدات الكهربائية بالأوم لمجتمع له وسط حسابى 90 أوم، وانحراف معيارى 2 أوم مبين فى شكل ٢ - ١٠.



شكل ١٠.٢: منحنى طبيعى للمقاومة الكهربائية لإحدى المعدات الكهربائية له وسط حسابى 90 أوم وانحراف معيارى 2 أوم

والكثير من التغيرات فى الطبيعة والصناعة تتبع التوزيع التكرارى للمنحنى الطبيعى . لهذا، التغيرات فى أوزان الأفيال، وسرعات الظباه، وأطوال الأفراد آدميين تتبع منحنى طبيعياً. كما أن التغيرات الموجودة فى الصناعة، مثل أوزان مسبوكات حديد الزهر، أو عمر اللمبات التى قوتها 60 وات، أو أبعاد حلقات مكبس صلب، يتوقع لها أن تتبع منحنى طبيعياً. عند اعتبار أطوال أفراد آدميين، يمكننا أن نتوقع أن تكون نسبة مئوية صغيرة منهم لها أطوال كبيرة جداً ونسبة صغيرة منهم لها أطوال قصيرة جداً، مع تقارب معظم الأطوال حول المتوسط. والمنحنى الطبيعى هو وصف جيد للتغيرات التى تحدث لمعظم خواص الجودة فى الصناعة التى تكون أساساً للعديد من أساليب مراقبة الجودة.

كل التوزيعات الطبيعية للمتغيرات المستمرة يمكن أن تحول إلى توزيع طبيعى نمطى (انظر شكل ٢ - ١٠) عن طريق استخدام قيمة طبيعية نمطية Standard-normal value . مثال ذلك، اعتبر القيمة 92 أوم فى شكل ٢ - ١٠، وهى عبارة عن انحراف معيارى واحد أكثر من الوسط الحسابى  $(\mu + 16 = 90 + 1(2))$  (92) . والتحويل إلى قيمة Z هو كما يلى :

$$Z = \frac{X_i - \mu}{\sigma} = \frac{92 - 90}{2} = +1$$

والذى يكون فوق الوسط الحسابى بمقدار انحراف معيارى واحد فى مقياس Z فى شكل ٢ - ١١ .

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} = 0.3989 e^{-z^2/2}$$

والصيغة للمنحنى الطبيعى النمطى هى :

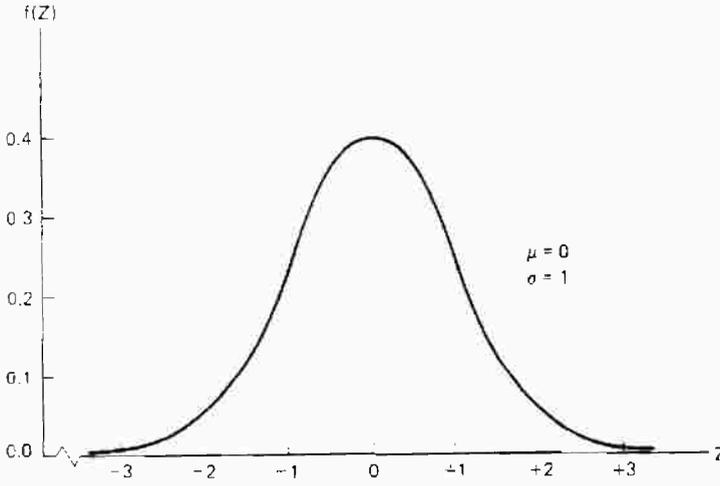
$$\text{حيث : } \pi = 3.142593$$

$$2.71828 = e$$

$$z = \frac{x_i - \mu}{\sigma}$$

يبين شكل ٢ - ١١ المنحنى النمطي بوسط حسابي صفر وانحراف معياري ١ .  
ويلاحظ أن المنحنى مقارب عند  $Z = -3$  و  $Z = 3$  .

والمساحة تحت المنحنى تساوي 1.000 أو 100% ولهذا يمكن أن تستخدم بسهولة في حسابات الاحتمالات. وحيث أن المساحة تحت المنحنى بين النقاط المختلفة تمثل إحصائيا مفيدا جدا، فيوجد جدول مساحة طبيعية في جدول أ في الملحق.



شكل ١١.٢: توزيع طبيعي نمطي بوسط حسابي صفر وانحراف معياري ١ .

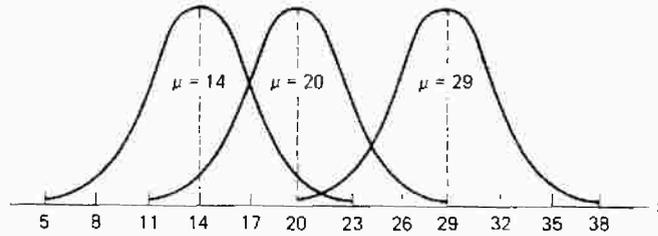
ويمكن الإشارة إلى التوزيع الطبيعي بأنه توزيع احتمالي طبيعي. وبينما يكون أكثر توزيعات المجتمعات أهمية، إلا أنه هناك عدد من التوزيعات الأخرى للمتغيرات المستمرة. فهناك أيضا عدد من التوزيعات الاحتمالية للمتغيرات الوثابة. وهذه التوزيعات تناقش في الفصل الرابع.

### العلاقة بالوسط الحسابي والانحراف المعياري

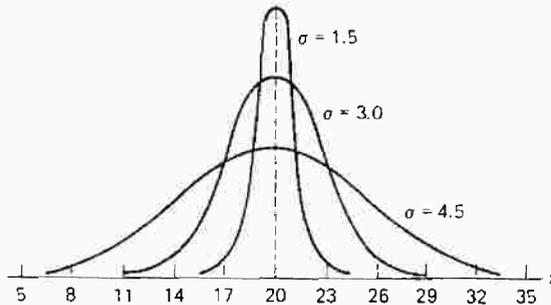
#### Relationship to the Mean and Standard Deviation

كما هو مبين بالصيغة الخاصة بالمنحنى الطبيعي النمطي، توجد علاقة محددة بين الوسط الحسابي، والانحراف المعياري، والمنحنى الطبيعي. يبين شكل ٢ - ١٢

ثلاثة منحنيات طبيعية لها قيم وسط حسابي مختلفة، ويلاحظ أن التغير الوحيد يوجد في الموقع فقط. وبين شكل ٢ - ١٣ ثلاثة منحنيات طبيعية لها نفس الوسط الحسابي ولكنها لها انحرافات معيارية مختلفة. ويوضح الشكل القاعدة بأنه كلما كبر الانحراف المعياري، كلما تسطح المنحنى (تكون البيانات مشتتة أكثر)، وكلما قل الانحراف المعياري، كلما ازدادت قمة المنحنى (يضيق تشتت البيانات). وإذا كان الانحراف المعياري صفراً، تكون كل القيم متطابقة مع الوسط الحسابي ولا يكون هناك منحنى على الإطلاق.



شكل ١٢.٢: منحني طبيعي بأوساط حسابية مختلفة وانحرافات معيارية متطابقة.

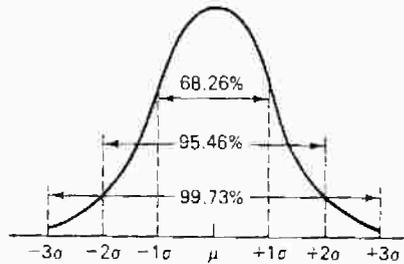


شكل ١٣.٢: منحني طبيعي بانحرافات معيارية مختلفة لكن بأوساط حسابية متطابقة.

يعرف المنحنى الطبيعي تعريفاً كاملاً بواسطة الوسط الحسابي والانحراف المعياري للمجتمع. وكذلك، كما هو مبين في شكل ٢ - ١٢ و ٢ - ١٣، هاتان المعلمتان تكونان مستقلتين. وفي كلمات أخرى، لا يؤثر التغيير في أحدهما على الآخر.

وتوجد علاقة بين الانحراف المعياري والمساحة تحت المنحنى كما هو مبين في شكل ٢ - ١٤. ويبين الشكل أنه في المنحنى الطبيعي 68.26% من العناصر تكون مشمولة بين حدى إنحراف معياري واحد ( $\mu + 1\sigma$  و  $\mu - 1\sigma$ )، وأن 95.46% من العناصر تكون مشمولة بين حدى انحرافين معيارين ( $\mu + 2\sigma$  و  $\mu - 2\sigma$ )، وأن 99.73% من العناصر تكون مشمولة بين حدى ثلاثة انحرافات معيارية.

( $\mu + 3\sigma$  و  $\mu - 3\sigma$ ) ويكون هناك 100% من العناصر مشمولة بين حدى الممالانهاية (الحد السالب والحد الموجب). وهذه النسب تكون حقيقية بغض النظر عن شكل المنحنى الطبيعي. والحقيقة أن 99.73% من العناصر تكون مشمولة بين حدى ثلاثة انحرافات معيارية هي أساس خرائط المراقبة التي تناقش في الفصل الثالث.



شكل ١٤.٢: النسب المئوية للعناصر المشمولة بين قيمتين معينتين للانحراف المعياري

## تطبيقات

## Applications

النسب المئوية للعناصر المشمولة بين أى قيمتين يمكن أن تحدد باستخدام حساب التفاضل والتكامل. إلا أن هذا ليس ضروريا، حيث أن المساحات تحت المنحنى لقيم  $Z$  المختلفة معطاه فى جدول أ فى الملحق. جدول أ، «المساحات تحت المنحنى الطبيعي» هو جدول قراءة من ناحية اليسار،<sup>(١)</sup> وهذا يعنى ان المساحات المعطاة تكون لجزء المنحنى من ناقص ما لانهاية إلى القيمة المحددة،  $X_i$ .

والخطوة الأولى هى تحديد قيمة  $Z$  باستخدام العلاقة :

$$Z = \frac{X_i - \mu}{\sigma}$$

حيث  $z$  = قيمة المنحنى النمطى

$x_i$  = القيمة الفردية

$\mu$  = الوسط الحسابى

$\sigma$  = الانحراف المعياري للمجتمع

بعد ذلك، باستخدام قيمة  $Z$  المحسوبة تحدد المساحة على يسار  $X_i$  من جدول أ. لهذا، إذا كانت  $Z$  المحسوبة  $-1.76$ ، تكون قيمة المساحة  $0.0392$ . وحيث أن المساحة الكلية تحت المنحنى هى  $1.000$ ، فإن قيمة  $0.0392$  من المساحة يمكن أن تتغير إلى نسبة مئوية للعناصر تحت المنحنى عن طريق نقل العلامة العشرية موقعين إلى اليمين. لهذا، يكون  $3.92\%$  من العناصر أقل من قيمة  $X_i$  معينة.

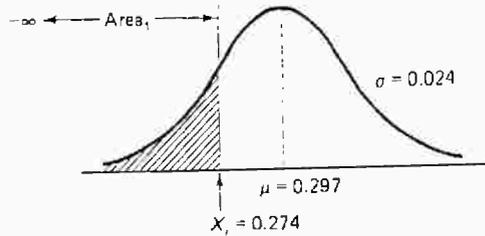
(١) فى بعض الكتب يرتب جدول المساحات تحت المنحنى الطبيعي بطريقة مختلفة

بافتراض ان البيانات موزعة توزيعا طبيعيا، يكون من الممكن إيجاد النسبة المئوية للعناصر فى البيانات التى تكون أقل من قيمة محددة، أو أكبر من قيمة محددة، أو بين قيمتين. وعندما تكون القيم مواصفات عليا و / أو دنيا، فتتوفر طريقة إحصائية بذلك قوية. ويوضح مثال المشكلة التالى هذه الطريقة .

### EXAMPLE PROBLEM

### مثال لمشكلة

قيمة الوسط الحسابى لوزن أحد أنواع الحبوب للسنة الماضية كانت 0.297 كج (10.5 أوقية) والانحراف المعياري كان 0.024 كج. بافتراض التوزيع الطبيعي، أوجد النسبة المئوية للبيانات التى تقع تحت ادنى حد مواصفات وهو 0.274 كج. (ملاحظة: حيث أن الوسط الحسابى والانحراف المعياري يتحددان من عدد كبير من الاختبارات على مدار السنة، فإنها تعتبر تقديرات صحيحة لقيم المجتمع).



$$Z = \frac{X_i - \mu}{\sigma}$$

$$= \frac{0.274 - 0.297}{0.024}$$

$$= -0.96$$

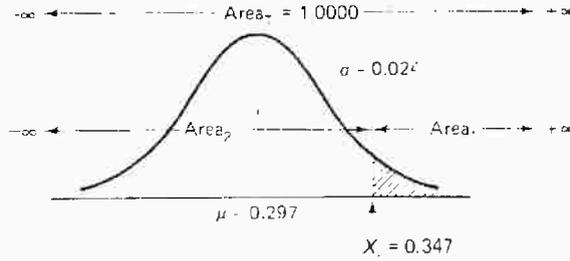
من جدول أ توجد قيمة Z وهي -0.96 ، والمساحة = 0.1685 أو 16.86% .  
لهذا، 16.85% من البيانات تكون أقل من 0.274

### EXAMPLE PROBLEM

### مثال لمشكلة

باستخدام بيانات المشكلة السابقة، حدد النسبة المئوية للبيانات التي تقع بعد 0.374 كج.

حيث أن جدول أ هو جدول قراءة من اليسار، فيطلب حل هذه المشكلة استخدام العلاقة :  $Area_1 + Area_2 = Area_T = 1.000$ . لهذا، تتحدد  $Area_2$  وتطرح من 1.000 للحصول على المساحة المطلوبة.



$$\begin{aligned} Z_2 &= \frac{X_1 - \mu}{\sigma} \\ &= \frac{0.347 - 0.297}{0.024} \\ &= + 2.08 \end{aligned}$$

من جدول أ يوجد أن لقيمة Z المساوية 2.08 ،  
 $Area_2 = 0.9812$

$$Area_1 = Area_T - Area_2 = 1.000 - 0.9812 = 0.018 \text{ أو } 1.88\%$$

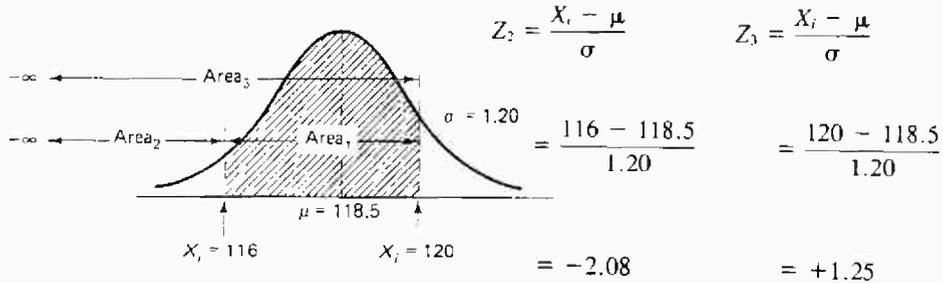
لهذا، 1.88% من البيانات تقع بعد 0.347 كج.

## EXAMPLE PROBLEM

## مثال لمشكلة

بين عدد كبير من اختبارات جهد الخط الكهربائي لساكنى أحد المنازل أن الوسط الحسابى 118.5 فولت وأن الانحراف المعياري للمجتمع 1.20 فولت. حدد النسبة المئوية للبيانات التى تقع بين 116 فولت و 120 فولت.

حيث أن جدول أ هو جدول قراءة من ناحية اليسار، فيتطلب الحل طرح المساحة الموجودة على يسار 116 فولت من المساحة الموجودة على يسار 120 فولت وبين الرسم والحسابات هذه الطريقة.



من جدول أ يوجد أنه لقيمة  $Z_2 = -2.08$ ، تكون  $Area_2 = 0.0188$  ولقيمة  $Z_3 = +1.20$ ، تكون  $Area_3 = 0.8944$ ، المساوية،

$$Area_1 = Area_3 - Area_2$$

$$87.56\% \text{ أو } = 0.8944 - 0.0188 = 0.8756$$

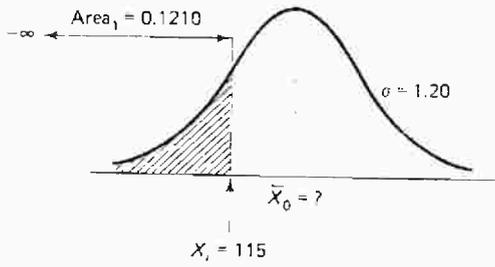
لهذا، 87.56% من البيانات تقع بين 116 فولت و 120 فولت.

## EXAMPLE PROBLEM

## مثال لمشكلة

من المرغوب فيه أن يكون 12.1% من جهد الخط الكهربائي أقل من 115 فولت، كيف يجب ضبط متوسط الجهد لتحقيق ذلك؟

حل هذا النوع من المشاكل يكون عكس المشاكل الأخرى. أولاً يوجد 12.10% أو 0.1210 في جسم الجدول أ. وهذا يعطى قيمة Z وباستخدام صيغة Z يمكننا أن نحل للحصول على الوسط الحسابى للفولت. من جدول أ، عند  $Area_1 = 0.1210$  تكون قيمة Z مساوية -1.17 والتي يتم الحصول عليها بالاستكمال.



$$Z_i = \frac{X_i - \bar{X}_0}{\sigma}$$

$$-1.17 = \frac{115 - \bar{X}_0}{1.20}$$

$$\bar{X}_0 = 116.4 \text{ V}$$

لهذا، يجب أن يتمركز الوسط الحسابى للفولت عند 116.4 فولت لتكون 12.1% من القيم أقل من 115 فولت.

لاحظ أنه تم التعويض عن  $\bar{x}_0$  في المعادلة. مفهوم المنحنى الطبيعي يكون مبنيًا على قيم  $\mu$  و  $\sigma$ ، إلا أن  $\bar{x}_0$  و  $s_0$ ، يمكن أن تعوض باشتراط أنه هناك إشارات بأن التوزيع طبيعي. ويوضح آخر مثال لمشكلة استقلالية  $\mu$  و  $\sigma$ . ولا يؤثر التغيير البسيط في مركزة العملية على التشتت.

## TESTS OF NORMALITY

## اختبارات الطبيعية

بسبب أهمية التوزيع الطبيعي، يتكرر ضرورة تحديد ما إذا كانت البيانات طبيعية

أم لا . وعند استخدام هذه الطريقة يجب أن يحتاط القارئ بأنه لا يوجد شيء أكيد %100 . طرق المدرج التكرارى، والانحراف، والتفرطح، ورسومات الاحتمالات، واختبار كاي تربيع تطبق أيضا مع بعض التعديل على توزيعات المجتمع الأخرى .

## Histogram

## المدرج التكرارى

الفحص البصرى لمدرج تكرارى يعد من كم كبير من البيانات يعطى مؤشرا لتوزيع المجتمع الأساسى . فإذا كان المدرج التكرارى أحادى المنوال، ومتماثل، وغير مستدق الطرف، فتكون الطبيعية إمكانية محددة ويمكن أن تكون معلومات كافية فى العديد من الحالات العملية . المدرج التكرارى فى شكل ٢ - ٤ لأوزان قضبان الصلب يكون أحادى المنوال، وغير مستدق الطرف، كما أنه متماثل بعض الشيء بالنسبة إلى الذيل العلوى . فإذا ما رفضت إحدى عمليات الفرز القضبان التى لها أوزان أكبر من 2.575، فإن هذا يمكن أن يوضح قطع الذيل العلوى .

وكلما ازداد حجم العينة، كلما تحسن الحكم على الطبيعية . ويوصى بحد أدنى لحجم العملية 50 .

## Skewness and Kurtosis

## الانحراف والتفرطح

مقاييس الانحراف والتفرطح هى اختبار آخر للطبيعية . من بيانات قضبان الصلب الموجودة فى جدول ٢ - ٦، نجد أن  $a_3 = -0.11$  وأن  $a_4 = 2.19$  . تحدد هذه القيم أن البيانات منحرفة قليلا ناحية اليسار إلا أنها قريبة من القيمة الطبيعية صفر، كما أن البيانات ليست لها قمة مثل التوزيع الطبيعى، والتى لها قيمة  $a_4 = 3.0$  .

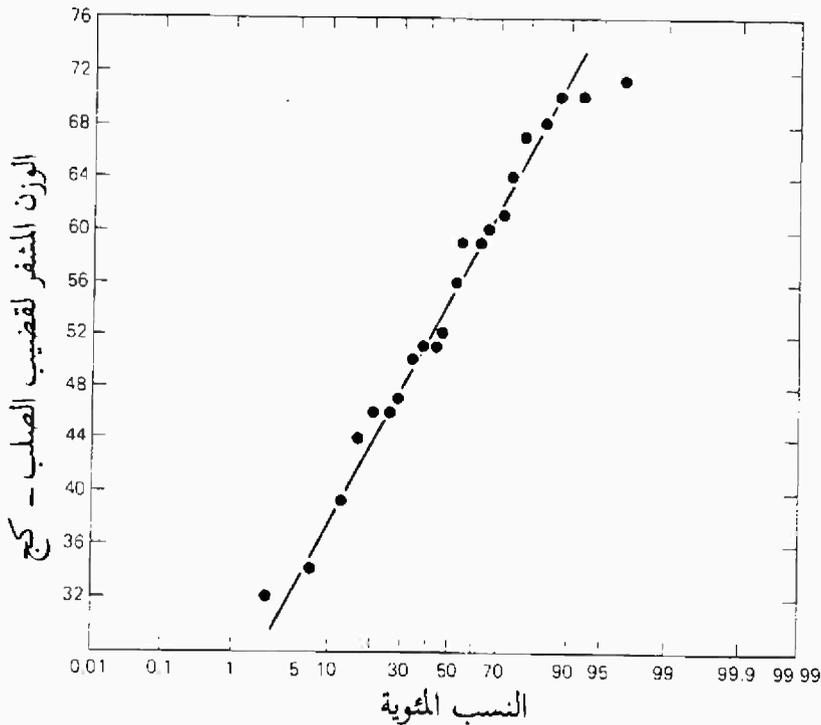
هذه القياسات تميل إلى أن تعطى نفس المعلومات مثل التى يقدمها المدرج التكرارى . وكما فى حالة المدرج التكرارى، كلما ازداد حجم العينة، كلما تحسن الحكم على الطبيعية . ويوصى بالقيمة 100 كحد أدنى لحجم العينة .

## Probability Plots

## رسومات الاحتمالات

اختبار آخر للطبيعية هو رسم البيانات على أوراق احتمالات طبيعية. هذا النوع من الأوراق مبين في شكل ٢ - ١٥. وتستخدم أوراق احتمالات مختلفة للتوزيعات المختلفة. لتوضيح الإجراء فإننا نستخدم بيانات قضبان الصلب مرة أخرى في صورتها المشفرة. وفيما يلي الإجراء خطوة بخطوة.

- ١- رتب البيانات order the data. البيانات من العمود الأول من جدول ٢ - ٤
- تستخدم في توضيح المفهوم. تسجل كل ملاحظة كما هو مبين في جدول ٢ - ١٣ من الأقل وحتى الأكبر. والملاحظات المزوجة تسجل كما هو مبين بواسطة القيمة 46.



شكل ١٥.٢: رسومات الاحتمالات للبيانات الموجودة في جدول ٢ - ١٣

## جدول ١٣.٢: بيانات عن أوزان قضبان الصلب لرسومات الاحتمالات

| الملاحظة<br>$x_i$ | الرتبة<br>$i$ | موقع الرسم | الملاحظة<br>$x_i$ | الرتبة<br>$i$ | موقع الرسم |
|-------------------|---------------|------------|-------------------|---------------|------------|
| 32                | 1             | 2.3        | 56                | 12            | 52.3       |
| 34                | 2             | 6.8        | 59                | 13            | 56.8       |
| 39                | 3             | 11.4       | 59                | 14            | 61.4       |
| 44                | 4             | 15.9       | 60                | 15            | 65.9       |
| 46                | 5             | 20.5       | 61                | 16            | 70.5       |
| 46                | 6             | 25.0       | 64                | 17            | 75.0       |
| 47                | 7             | 29.5       | 67                | 18            | 79.5       |
| 50                | 8             | 34.1       | 68                | 19            | 84.1       |
| 51                | 9             | 38.6       | 70                | 20            | 88.6       |
| 51                | 10            | 43.2       | 70                | 21            | 93.2       |
| 52                | 11            | 47.7       | 71                | 22            | 97.7       |

٢- أعط رتب للملاحظات rank the observations. بدءا بالرقم 1 لأقل ملاحظة، و 2 للملاحظة التالية، وهكذا، أعط رتبا للملاحظات. والرتب مبينة في العمود الثاني من جدول ٢ - ١٣.

٣- احسب موقع الرسم calculate the plotting position. تتحقق هذه الخطوة باستخدام الصيغة :

$$PP = \frac{100(i - 0.5)}{n}$$

حيث :  $i$  = الرتبة

$PP$  = موقع الرسم كنسبة مئوية

$n$  = حجم العينة

أول موقع رسم هو  $2.3 = 100(1 - 0.5) / 22$  ، أى 2.3% ، وتحسب المواقع الأخرى بنفس الطريقة وتوضع في جدول ٢ - ١٣.

٤- سَمِّ مقياس البيانات label the data scale . تتراوح قيم البيانات المشفرة من 32 إلى 71، لهذا فالمقياس الرأسى يسمى بصورة مناسبة ويبين فى شكل ٢ - ١٥ . ويمثل المقياس الأفقى المنحنى الطبيعى ويكون مطبوعا بصورة مسبقة على الأوراق .

٥- ارسم النقاط plot the points . يرسم موقع الرسم والملاحظة على ورق المنحنى الطبيعى .

٦- حاول أن تضبط بالعين «أفضل» خط "best" attempt to fit by eye a line . المسطرة البلاستيكية الشفافة تكون أكثر فائدة فى عمل هذا الحكم . عند ضبط هذا الخط، يجب أن تعطى أوزانا أكبر للقيم المركزية عن القيم الكبيرة جدا أو الصغيرة جدا .

٧- حدد الطبيعية determine normality . هذا القرار هو أحد الأحكام مثل درجة قرب النقاط إلى الخط المستقيم . إذا ما ألغينا النقاط الكبيرة جدا أو الصغيرة جدا عند طرفى الخط، فيمكننا أن نفترض بصورة معقولة أن البيانات موزعة توزيعا طبيعيا .

إذا بدت الطبيعية معقولة، فيمكن الحصول على معلومات إضافية من الرسم . يقع الوسط الحسابى عند النسبة المئوية الخمسون، والذى يعطى قيمة تقريبية 55 . الانحراف المعيارى هو خمسان الفرق بين النسبة المئوية التسعين والنسبة المئوية العاشرة، والتى يمكن أن تكون [38 - 72) (2/5) 14] . كما يمكننا أن نستخدم الرسم أيضا فى تحديد النسبة المئوية للبيانات الأقل من أو الأكبر من أو بين قيم بيانات محددة . مثال ذلك، النسبة المئوية الأقل من 48 هى بالتقريب 31% . وبالرغم من أن مثال المشكلة يستخدم 22 نقطة بيانات مع الحصول على نتائج جيدة، فإن أقل حجم عينة يوصى به هو 30 .

## Chi-Square

## كاي تربيع

اختبار كاي تربيع هو طريقة أخرى لتحديد ما إذا كانت بيانات العينة تناسب توزيعاً طبيعياً أو أى توزيع آخر. وهذه الطريقة تقع خارج مدى هذا الكتاب، إلا أنها متاحة فى معظم كتب الإحصاء. ويوصى بأن يكون حجم العينة 125 على الأقل.

ومن المهم للمحلل أن يفهم أن أى طريقة من هذه الطرق لا تثبت أن البيانات موزعة توزيعاً طبيعياً. وإنما يمكننا أن نستخلص فقط أنه لا يوجد دلائل على أن البيانات لا يمكن أن تعامل كما لو كانت موزعة توزيعاً طبيعياً.

## COMPUTER PROGRAM

## برنامج حاسوب

توجد عينات برامج حاسوب لهذا الفصل والفصول الثالث، والرابع، والخامس، والسادس، والسابع. وقد صممت هذه البرامج للعمل على أجهزة حاسوب شخصية وكتبت ببيسك ميكروسوفت واختبرت بجهاز حاسوب شخصى من طراز-Zenith Z-89. وقد قدمت البرامج فى صورة بسيطة بقدر الإمكان وتتبع نصوص الكتاب بحيث يستطيع القارئ أن يفهم البرمجة. والاحتياطات لفرز بيانات المدخلات وتصحيح أخطاء بيانات المدخلات ليست مشمولة فى البرامج.

وعينة البرنامج الموجودة فى هذا الفصل تحسب الإحصائيات الرئيسية السبع وتحدد التوزيع التكرارى. وهى مبينة فى شكل ٢ - ١٦. وإذا ما كان هناك رغبة فى طباعة البيانات، فيجب استخدام عبارة LPRINT بعد العبارة رقم 110. كذلك، يمكن استخدام عبارة شفرة بعد العبارة رقم 110. وقد أخذ الاحتياط فى البرنامج لتغيير التوزيع التكرارى عن طريق تغيير فترة الخلية و/أو أقل قيمة مركز للخلية. والبيانات المستخدمة مع البرنامج هى الموجودة فى جدول ٢ - ٤.

ويمكن تعزيز البرنامج عن طريق رسم التوزيع التكرارى على هيئة مدرج تكرارى. وهذا النشاط يعتمد على وحدة مخرجات الرسومات المتاحة للقارئ. وكذلك، يمكن إعداد التكرار النسبى، والتكرار المتجمع الصاعد، والتكرار النسبى المتجمع الصاعد بإضافة بضع عبارات إضافية.

```

10 REM                               STAT-PACK
20 REM
30 REM           Average, Median, Range, Standard Deviation
40 REM           Variance, Kurtosis, Skewness,
50 REM           Frequency Distribution
60 REM
70 DIM X(600), MP(20), LB(20), UB(20), F(600)
80 PRINT " Enter number of data points." : INPUT N
90 PRINT " Enter data."
100      FOR I = 1 TO N
110      INPUT X(I)
120      NEXT I
130 REM                               Average
140 SX = 0
150      FOR I=1 TO N
160      SX = SX + X(I)
170      NEXT I
180 AVG = SX / N
190 LPRINT TAB(5); " Average = ";AVG
200 REM                               Sample Standard Deviation, Variance,
210 REM                               Skewness, Kurtosis
220 S2 = 0 : S3 = 0 : S4 = 0
230      FOR I = 1 TO N
240      D = X(I) - AVG
250      S2 = S2 + D^2
260      S3 = S3 + D^3
270      S4 = S4 + D^4
280      NEXT I
290 SD = SQR(S2 / (N-1))
300 VA = SD^2
310 A3 = S3 / N / SD^3
320 A4 = S4 / N / SD^4
330 LPRINT TAB(5); " Sample Standard Deviation = ";SD
340 LPRINT TAB(5); " Variance = ";VA
350 LPRINT TAB(5); " Skewness = ";A3
360 LPRINT TAB(5); " Kurtosis = ";A4
370 REM                               Sort Routine
380      FOR J = 1 TO (N - 1)
390      K = N - J
400      FOR I = 1 TO K
410      Q = I + 1

```

شكل ١٦.٢: برنامج حاسوب بلغة البيسك لحساب الإحصائيات السبع الرئيسية والتوزيع التكرارى

```

420     IF X(1) < X(Q) GOTO 460
430     A = X(I)
440     X(I) = X(Q)
450     X(Q) = A
460     NEXT I
470     NEXT J
480 REM                                     Range
490 R=X(N)-X(1)
500 LPRINT TAB(5);" Range = ";R
510 REM                                     Median
520 J = (N + 1) / 2
530 K = (N + 1) / 2
540 MD = (X(J) + X(K)) / 2
550 LPRINT TAB(5);" Median = "; MD
560 REM                                     Frequency Distribution
570 PRINT " Enter Interval---odd value preferred.
580 INPUT IN
590 PRINT " Enter 1 if lowest cell midpoint is lowest X(I) value."
600 PRINT " Enter 0 if another lowest cell midpoint is desired."

610 INPUT M
620 IF M = 1 THEN MP(1) = X(1) : GOTO 650
630 PRINT " Enter midpoint value of lowest cell."
640 INPUT MP(1)
650 LB(1) = MP(1) - IN / 2
660 UB(1) = MP(1) + IN / 2
670 MPS = MP(1) : LBS = LB(1) : UBS = UB(1)
680 FOR K = 2 TO 20
690     MPS = MPS + IN
700     MP(K) = MPS
710     LBS = LBS + IN
720     LB(K) = LBS
730     UBS = UBS + IN
740     UB(K) = UBS
750     IF UB(K) >= X(N) GOTO 810
760     NEXT K
770 IF UB(K - 1) < X(N) GOTO 790
780 GOTO 810
790 PRINT " Cell interval too small select a larger one."
800 GOTO 570
810 H=K
820 FOR J=1 TO H
830     FS = 0
840     FOR I = 1 TO N
850         IF X(I) < LB(J) GOTO 890
860         IF X(I) > UB(J) GOTO 890
870         FS = FS + 1
880     F(J) = FS
890     NEXT I
900     NEXT J
910 LPRINT : LPRINT
920 LPRINT TAB(10);" BOUNDARIES";TAB(29);" MIDPOINT";TAB(45);" FREQUENCY"
930     FOR J = 1 TO H
940         LPRINT TAB(7); LB(J);"/";UB(J);TAB(31); MP(J);TAB(49); F(J)
950     NEXT J
960 PRINT " If the frequency distribution is satisfactory, Enter 1."
970 PRINT " If the frequency distribution is unsatisfactory, Enter 0."
980 PRINT " and select another interval and/or midpoint."
990 INPUT M
1000 IF M = 1 GOTO 1020
1010 IF M = 0 GOTO 570
1020 END

```

Average = 2.554  
 Sample Standard Deviation = .0111273  
 Variance = 1.23817E-04  
 Skewness = -.0837673  
 Kurtosis = 2.22477  
 Range = .0440002  
 Median = 2.553

| BOUNDARIES      | MIDPOINT | FREQUENCY |
|-----------------|----------|-----------|
| 2.5305 / 2.5355 | 2.533    | 6         |
| 2.5355 / 2.5405 | 2.538    | 8         |
| 2.5405 / 2.5455 | 2.543    | 12        |
| 2.5455 / 2.5505 | 2.548    | 13        |
| 2.5505 / 2.5555 | 2.553    | 20        |
| 2.5555 / 2.5605 | 2.558    | 19        |
| 2.5605 / 2.5655 | 2.563    | 13        |
| 2.5655 / 2.5705 | 2.568    | 11        |
| 2.5705 / 2.5755 | 2.573    | 8         |

## PROBLEM

## مشاكل

١- قرب الأعداد التالية إلى أقرب موقعين عشريين :

أ- 0.862      ب- 0.625      ج- 0.149      د- 0.475

٢- نفذ العملية المحددة مع إعطاء الإجابة بالأعداد الصحيحة للأرقام المنوية.

أ- (8.20) (34.6)      ب- (635) (0.035)      ج- 3.8735/6.1  
 د- 5.362/6 (6 هو عدد للعد).

٣- نفذ العملية المحددة مع إعطاء الإجابة بالأعداد الصحيحة للأرقام المنوية.

أ-  $64.3 + 2.05$       ب-  $383.0 - 1.95$       ج-  $4 - 8.652$  (4 ليس عددا للعد)  
 د-  $4 - 8.652$  (4 عدد للعد)      هـ-  $24.32 - 6.4 \times 10^2$

٤- لقد سجل أحد لاعبي كرة السلة في آخر 70 مباراة له البيانات التالية :

|    |    |    |    |    |    |    |
|----|----|----|----|----|----|----|
| 10 | 17 | 9  | 17 | 18 | 20 | 16 |
| 7  | 17 | 19 | 13 | 15 | 14 | 13 |
| 12 | 13 | 15 | 14 | 13 | 10 | 14 |
| 11 | 15 | 14 | 11 | 15 | 15 | 16 |
| 9  | 18 | 15 | 12 | 14 | 13 | 14 |
| 13 | 14 | 16 | 15 | 16 | 15 | 15 |
| 14 | 15 | 15 | 16 | 13 | 12 | 16 |
| 10 | 16 | 14 | 13 | 16 | 14 | 15 |
| 6  | 15 | 13 | 16 | 15 | 16 | 16 |
| 12 | 14 | 16 | 15 | 16 | 13 | 15 |

أ- اعمل استمارة (ورقة) عد في ترتيب تصاعدي

ب - مستخدما البيانات السابقة، إرسم مدرجا تكراريا.

٥- تملأ إحدى الشركات زجاجات من الشامبو وتحاول أن تحتفظ بوزن محدد لمنتجها. ويقدم الجدول أوزان 110 زجاجة من زجاجات الشامبو التي اختبرت على فترات عشوائية. اعمل عدا لهذه الأوزان وارسم مدرجا تكراريا لها. (الأوزان بالكيلوجرام).

|      |      |      |      |      |      |      |      |      |      |      |
|------|------|------|------|------|------|------|------|------|------|------|
| 6.00 | 5.98 | 6.01 | 6.01 | 5.97 | 5.99 | 5.98 | 6.01 | 5.99 | 5.98 | 5.96 |
| 5.98 | 5.99 | 5.99 | 6.03 | 5.99 | 6.01 | 5.98 | 5.99 | 5.97 | 6.01 | 5.98 |
| 5.97 | 6.01 | 6.00 | 5.96 | 6.00 | 5.97 | 5.95 | 5.99 | 5.99 | 6.01 | 6.00 |
| 6.01 | 6.03 | 6.01 | 5.99 | 5.99 | 6.02 | 6.00 | 5.98 | 6.01 | 5.98 | 5.99 |
| 6.00 | 5.98 | 6.05 | 6.00 | 6.00 | 5.98 | 5.99 | 6.00 | 5.97 | 6.00 | 6.00 |
| 6.00 | 5.98 | 6.00 | 5.94 | 5.99 | 6.02 | 6.00 | 5.98 | 6.02 | 6.01 | 6.00 |
| 5.97 | 6.01 | 6.04 | 6.02 | 6.01 | 5.97 | 5.99 | 6.02 | 5.99 | 6.02 | 5.99 |
| 6.02 | 5.99 | 6.01 | 5.98 | 5.99 | 6.00 | 6.02 | 5.99 | 6.02 | 5.95 | 6.02 |
| 5.96 | 5.99 | 6.00 | 6.00 | 6.01 | 5.99 | 5.96 | 6.01 | 6.00 | 6.01 | 5.98 |
| 6.00 | 5.99 | 5.98 | 5.99 | 6.03 | 5.99 | 6.02 | 5.98 | 6.02 | 6.02 | 5.97 |

٦- فيما يلي 125 قراءة تم الحصول عليها من إحدى المستشفيات بواسطة محلل لإحدى دراسات الوقت والحركة الذي أخذ خمس قراءات كل يوم لمدة 25 يوما. اعمل ورقة للعد. ارسم جدولاً مبيّناً مراكز الخلايا، والتكرارات الملاحظة. ارسم مدرجاً تكرارياً.

| اليوم | استمرارية وقت العملية<br>دقيقة |      |      |      |      |
|-------|--------------------------------|------|------|------|------|
| 1     | 1.90                           | 1.93 | 1.95 | 2.05 | 2.20 |
| 2     | 1.76                           | 1.81 | 1.81 | 1.83 | 2.01 |
| 3     | 1.80                           | 1.87 | 1.95 | 1.97 | 2.07 |
| 4     | 1.77                           | 1.83 | 1.87 | 1.90 | 1.93 |
| 5     | 1.93                           | 1.95 | 2.03 | 2.05 | 2.14 |
| 6     | 1.76                           | 1.88 | 1.95 | 1.97 | 2.00 |
| 7     | 1.87                           | 2.00 | 2.00 | 2.03 | 2.10 |
| 8     | 1.91                           | 1.92 | 1.94 | 1.97 | 2.05 |
| 9     | 1.90                           | 1.91 | 1.95 | 2.01 | 2.05 |
| 10    | 1.79                           | 1.91 | 1.93 | 1.94 | 2.10 |
| 11    | 1.90                           | 1.97 | 2.00 | 2.06 | 2.28 |
| 12    | 1.80                           | 1.82 | 1.89 | 1.91 | 1.99 |
| 13    | 1.75                           | 1.83 | 1.92 | 1.95 | 2.04 |
| 14    | 1.87                           | 1.90 | 1.98 | 2.00 | 2.08 |
| 15    | 1.90                           | 1.95 | 1.95 | 1.97 | 2.03 |
| 16    | 1.82                           | 1.99 | 2.01 | 2.06 | 2.06 |
| 17    | 1.90                           | 1.95 | 1.95 | 2.00 | 2.10 |
| 18    | 1.81                           | 1.90 | 1.94 | 1.97 | 1.99 |
| 19    | 1.87                           | 1.89 | 1.98 | 2.01 | 2.15 |
| 20    | 1.72                           | 1.78 | 1.96 | 2.00 | 2.05 |
| 21    | 1.87                           | 1.89 | 1.91 | 1.91 | 2.00 |
| 22    | 1.76                           | 1.80 | 1.91 | 2.06 | 2.12 |
| 23    | 1.95                           | 1.96 | 1.97 | 2.00 | 2.00 |
| 24    | 1.92                           | 1.94 | 1.97 | 1.99 | 2.00 |
| 25    | 1.85                           | 1.90 | 1.90 | 1.92 | 1.92 |

٧- اختبرت القوة النسبية لعدد 150 سبيكة لحام فضية، والنتائج معطاة في الجدول التالي. عد هذه الأرقام ورتبها في توزيع تكرارى. حدد فترة الخلية والعدد التقريبي للخلايا. ارسم جدولاً مبيّناً مراكز الخلايا، وحدود الخلايا، والتكرارات الملاحظة. ارسم المدرج التكرارى.

|     |     |     |     |     |     |
|-----|-----|-----|-----|-----|-----|
| 1.5 | 1.2 | 3.1 | 1.3 | 0.7 | 1.3 |
| 0.1 | 2.9 | 1.0 | 1.3 | 2.6 | 1.7 |
| 0.3 | 0.7 | 2.4 | 1.5 | 0.7 | 2.1 |
| 3.5 | 1.1 | 0.7 | 0.5 | 1.6 | 1.4 |
| 1.7 | 3.2 | 3.0 | 1.7 | 2.8 | 2.2 |
| 1.8 | 2.3 | 3.3 | 3.1 | 3.3 | 2.9 |
| 2.2 | 1.2 | 1.3 | 1.4 | 2.3 | 2.5 |
| 3.1 | 2.1 | 3.5 | 1.4 | 2.8 | 2.8 |
| 1.5 | 1.9 | 2.0 | 3.0 | 0.9 | 3.1 |
| 1.9 | 1.7 | 1.5 | 3.0 | 2.6 | 1.0 |
| 2.9 | 1.8 | 1.4 | 1.4 | 3.3 | 2.4 |
| 1.8 | 2.1 | 1.6 | 0.9 | 2.1 | 1.5 |
| 0.9 | 2.9 | 2.5 | 1.6 | 1.2 | 2.4 |
| 3.4 | 1.3 | 1.7 | 2.6 | 1.1 | 0.8 |
| 1.0 | 1.5 | 2.2 | 3.0 | 2.0 | 1.8 |
| 2.9 | 2.5 | 2.0 | 3.0 | 1.5 | 1.3 |
| 2.2 | 1.0 | 1.7 | 3.1 | 2.7 | 2.3 |
| 0.6 | 2.0 | 1.4 | 3.3 | 2.2 | 2.9 |
| 1.6 | 2.3 | 3.3 | 2.0 | 1.6 | 2.7 |
| 1.9 | 2.1 | 3.4 | 1.5 | 0.8 | 2.2 |
| 1.8 | 2.4 | 1.2 | 3.7 | 1.3 | 2.1 |
| 2.9 | 3.0 | 2.1 | 1.8 | 1.1 | 1.4 |
| 2.8 | 1.8 | 1.8 | 2.4 | 2.3 | 2.2 |
| 2.1 | 1.2 | 1.4 | 1.6 | 2.4 | 2.1 |
| 2.0 | 1.1 | 3.8 | 1.3 | 1.3 | 1.0 |

٨ - باستخدام بيانات المشكلة رقم ٤ ، ارسم :

أ - مدرجا تكراريا نسبيا .

ب - مدرجا تكراريا متجمعا صاعدا .

ج - مدرجا تكراريا متجمعا صاعدا نسبيا .

٩ - باستخدام بيانات المشكلة رقم ٥ ، ارسم :

أ - مدرجا تكراريا نسبيا .

ب - مدرجا تكراريا متجمعا صاعدا .

ج - مدرجا تكراريا متجمعا صاعدا نسبيا .

١٠- باستخدام بيانات المشكلة رقم ٦ ، ارسم :

أ - مدرجا تكراريا نسبيا .

ب - مدرجا تكراريا متجمعا صاعدا .

ج - مدرجا تكراريا متجمعا صاعدا نسبيا .

١١- باستخدام بيانات المشكلة رقم ٧ ، ارسم :

أ - مدرجا تكراريا نسبيا .

ب - مدرجا تكراريا متجمعا صاعدا .

ج - مدرجا تكراريا متجمعا صاعدا نسبيا .

١٢- ارسم خريطة قضبان للبيانات الموجودة فى :

أ - المشكلة رقم ٤ .

ب - المشكلة رقم ٥ .

١٣- باستخدام بيانات المشكلة رقم ٦ ، ارسم :

أ - مضلعا تكراريا .

ب - منحنى متجمع صاعد .

١٤- باستخدام بيانات المشكلة رقم ٧ ، ارسم :

أ - مضلعا تكراريا .

ب - منحنى متجمع صاعد .

١٥- يختبر أحد التقنيين الكهربائيين شدة التيار الداخلى لأحد المنازل السكنية

وحصل على الخمس قراءات التالية : 115 و 113 و 121 و 115 و 116 . ما

المتوسط ؟

١٦- قام أحد العاملين بشمان رحلات لتحصيل مقطورة. إذا كانت مسافات الرحلات بالمترو كانت كما يلي : 25.6 و 24.8 و 22.6 و 21.3 و 19.6 و 18.5 و 16.2 و 15.5، ما المتوسط؟

١٧- تعطى معدلات الضوضاء فى مواقع محددة فى ورشة عمل أختام فى التوزيع التكرارى التالى. وتقاس الضوضاء بالديسيبل. حدد المتوسط؟

| التكرار | نقطة مركز الخلية |
|---------|------------------|
| 2       | 148              |
| 3       | 139              |
| 8       | 130              |
| 11      | 121              |
| 27      | 112              |
| 35      | 103              |
| 43      | 94               |
| 33      | 85               |
| 20      | 76               |
| 12      | 67               |
| 6       | 58               |
| 4       | 49               |
| 2       | 40               |

١٨- كانت أوزان 65 من المسبوكات بالكيلوجرام موزعة على النحو التالى :

| التكرار | نقطة مركز الخلية |
|---------|------------------|
| 6       | 3.5              |
| 9       | 3.8              |
| 18      | 4.1              |
| 14      | 4.4              |
| 13      | 4.7              |
| 5       | 5.0              |

حدد المتوسط ..

١٩- الاختبارات التدميرية على عمر مكونات اليكترونية أجريت في حالتين مختلفتين. في الحالة الأولى كان لثلاثة اختبارات وسطا حسابيا مقداره 3320 ساعة، وفي الحالة الثانية كان لاختبارين وسطا حسابيا مقداره 3180 ساعة. ما المتوسط المرجح؟

٢٠- متوسط طول 24 طالبا في القسم رقم 1 لمقرر مراقبة الجودة كان 1.75 مترا، ومتوسط طول 18 طالبا في القسم رقم 2 لمقرر مراقبة الجودة كان 1.79 مترا، ومتوسط طول 29 طالبا في القسم رقم 3 لمقرر مراقبة الجودة كان 1.68 مترا. ما متوسط طول الطلبة في الثلاثة أقسام لمقرر مراقبة الجودة؟

٢١- حدد الوسيط للأعداد التالية :

أ - 22 و 11 و 15 و 8 و 18.

ب - 35 و 28 و 33 و 38 و 43 و 36 .

٢٢- حدد الوسيط لكل مما يلي :

أ - التوزيع التكرارى للمشكلة رقم ١٧ .

ب - التوزيع التكرارى للمشكلة رقم ١٨ .

ج - التوزيع التكرارى للمشكلة رقم ٢٨ .

د - التوزيع التكرارى للمشكلة رقم ٣٠ .

هـ - التوزيع التكرارى للمشكلة رقم ٦ .

و - التوزيع التكرارى للمشكلة رقم ٧ .

٢٣- بمعرفة السلسلة التالية من الأعداد، حدد المنوال.

أ - 50 و 45 و 55 و 55 و 45 و 50 و 55 و 45 و 55.

ب - 89 و 87 و 88 و 83 و 86 و 82 و 84 .

ج - 11 و 17 و 14 و 12 و 12 و 14 و 14 و 15 و 17 و 17 .

٢٤- حدد خلية المنوال للبيانات فى كل مما يلى :

أ - المشكلة رقم ٤ .

ب - المشكلة رقم ٥ .

ج - المشكلة رقم ٦ .

د - المشكلة رقم ٧ .

هـ - المشكلة رقم ١٧ .

و - المشكلة رقم ١٨ .

٢٥- حدد المدى لكل فئة من فئات الأعداد التالية :

أ - 16 و 25 و 18 و 17 و 16 و 21 و 14

ب - 45 و 39 و 42 و 42 و 43

ج - بيانات المشكلة رقم ٤

د - بيانات المشكلة رقم ٥

٢٦- اختبارات الترددات على قضبان نحاسية طولها 145 سم أعطت القيم

التالية : 1200 و 1190 و 1205 و 1185 و 1200 ذبذبة فى الثانية. ما الانحراف

المعيارى للعينة؟

٢٧- كانت أربع قراءات لسلك الورق فى هذا الكتاب كما يلى : 0.076 مم و

0.082 مم و 0.073 مم و 0.077 مم. حدد الانحراف المعيارى للعينة؟.

٢٨- يبين التوزيع التكرارى المعطى هنا النسبة المئوية للكبريت العضوى فى أحد أنواع الفحم. حدد الانحراف المعيارى للعينة؟.

| نقطة مركز الخلية<br>(%) | التكرار<br>(عدد العينات) |
|-------------------------|--------------------------|
| 0.5                     | 1                        |
| 0.8                     | 16                       |
| 1.1                     | 12                       |
| 1.4                     | 10                       |
| 1.7                     | 12                       |
| 2.1                     | 18                       |
| 2.4                     | 16                       |
| 2.7                     | 3                        |

٢٩- حدد الانحراف المعيارى للعينة لكل مما يلى :

أ- بيانات المشكلة رقم ٧.

ب- بيانات المشكلة رقم ١٧.

٣٠- حدد المتوسط والانحراف المعيارى للعينة للتوزيع التكرارى لعدد

الفحوصات فى اليوم الواحد التى تلى :

| نقطة مركز الخلية | التكرار |
|------------------|---------|
| 1000             | 6       |
| 1300             | 13      |
| 1600             | 22      |
| 1900             | 17      |
| 2200             | 11      |
| 2500             | 8       |

٣١- باستخدام بيانات المشكلة رقم ١٧ ، ارسم :

أ- مضلع تكرارى.

ب- منحنى متجمعا صاعداً.

٣٢- باستخدام بيانات المشكلة رقم ١٨ ، ارسم :  
أ - مضلعا تكراريا.

ب - منحى متجمعا صاعدا.

٣٣- باستخدام بيانات المشكلة رقم ٢٨ ، ارسم :  
أ - مضلعا تكراريا.

ب - منحى متجمعا صاعدا.

٣٤- باستخدام بيانات المشكلة رقم ٣٠ ، ارسم :  
أ - مضلعا تكراريا.

ب - منحى متجمعا صاعدا.

٣٥- باستخدام بيانات المشكلة رقم ١٧ ، ارسم :  
أ - مدرجا تكراريا.

ب - مدرجا تكراريا نسبيا.

ج - مدرجا تكراريا متجمعا صاعدا.

د - مدرجا تكراريا متجمعا صاعدا نسبيا.

٣٦- باستخدام بيانات المشكلة رقم ١٨ ، ارسم :  
أ - مدرجا تكراريا.

ب - مدرجا تكراريا نسبيا.

ج - مدرجا تكراريا متجمعا صاعدا.

د - مدرجا تكراريا متجمعا صاعدا نسبيا.

٣٧- باستخدام بيانات المشكلة رقم ٢٨ ، ارسم :

أ - مدرجا تكراريا .

ب - مدرجا تكراريا نسبيا .

ج - مدرجا تكراريا متجمعا صاعدا .

د - مدرجا تكراريا متجمعا صاعدا نسبيا .

٣٨- باستخدام بيانات المشكلة رقم ٣٠ ، ارسم :

أ - مدرجا تكراريا .

ب - مدرجا تكراريا نسبيا .

ج - مدرجا تكراريا متجمعا صاعدا .

د - مدرجا تكراريا متجمعا صاعدا نسبيا .

٣٩- حدد الالتواء والتفرطح لكل مما يلي :

أ - المشكلة رقم ٤ .

ب - المشكلة رقم ٥ .

ج - المشكلة رقم ٦ .

د - المشكلة رقم ٧ .

هـ - المشكلة رقم ١٨ .

و - المشكلة رقم ٣٠ .

٤٠- إذا كانت أقصى ضوضاء مسموح بها هي 134.5 ديسيبل، ما النسبة

المقوية لبيانات المشكلة رقم ١٧ التي تقع أعلى من هذه القيمة؟

٤١- قوم المدرج التكرارى للمشكلة رقم ١٨ ، حيث تكون المواصفات هي:  
 $4.25 + 0.60$  كج.

٤٢- لا تستخدم إحدى شركات المنافع فحما إذا ما احتوى على أكثر من  
 2.25% كبريت. اعتمادا على المدرج التكرارى الموجود فى المشكلة رقم ٢٨ ، ما  
 النسبة المئوية للفحم التى تقع فى هذه الفئة؟

٤٣- الوسط الحسابى للمجتمع لشركة دراجات سباق كان 9.07 كج (20.0  
 رطل) مع انحراف معيارى للمجتمع مقداره 40.0 كج. إذا كان التوزيع طبيعيا  
 تقريبا، حدد (أ) النسبة المئوية للدراجات التى تقل عن 8.30 كج، (ب) النسبة  
 المئوية للدراجات التى تزيد عن 10.00 كج، و (ج) النسبة المئوية للدراجات التى  
 تقع بين 8.00 و 10.10 كج.

٤٤- إذا كان الوسط الحسابى للوقت اللازم لتنظيف الغرفة بأحد الفنادق هو  
 16.0 دقيقة والانحراف المعيارى هو 1.5 دقيقة، ما النسبة المئوية للغرف التى تأخذ  
 أقل من 13.0 دقيقة فى تنظيفها؟ وما النسبة المئوية للغرف التى تأخذ من 13.0 إلى  
 20.5 دقيقة فى تنظيفها؟ البيانات تتبع توزيعا طبيعيا.

٤٥- أحد مصنعى الحبوب الباردة يريد أن يكون 1.5% من المنتج أقل من  
 مواصفات الوزن وهى 0.567 كج (1.25 رطل). إذا كانت البيانات تتبع توزيعا  
 طبيعيا وكان الانحراف المعيارى لآلة ملء الحبوب هو 0.018 كج، ما الوسط  
 الحسابى اللازم للوزن؟

٤٦- فى التجليخ المتقن لأحد الأجزاء المعقدة، من الأكثر اقتصادا إعادة تشغيل  
 الجزء عن إهماله كخردة. لهذا، من المرغوب فيه تحديد النسبة المئوية لإعادة  
 التشغيل عند 12.5%. بافتراض أن البيانات تتبع توزيعا طبيعيا، وأن الانحراف  
 المعيارى كان 0.01 مم، والحد العلوى للمواصفات كان 25.38 مم (0.99 بوصة)،  
 حدد مركز العملية.

٤٧- باستخدام معلومات المشكلة رقم ٣٩، ما هو حكمك الخاص بالطبيعية للتوزيع؟

أ- الخاص بالمشكلة رقم ٤ .

ب- الخاص بالمشكلة رقم ٥ .

ج- الخاص بالمشكلة رقم ٦ .

د- الخاص بالمشكلة رقم ٧ .

هـ- الخاص بالمشكلة رقم ١٨ .

و- الخاص بالمشكلة رقم ٣٠ .

٤٨- باستخدام ورقة الاحتمالات الطبيعية، حدد (احكم على) الطبيعية لتوزيع كل مما يلي :

أ- العمود الثاني لجدول ٢ - ٤ .

ب- أول ثلاثة أعمدة للمشكلة رقم ٥ .

ج- العمود الثاني للمشكلة رقم ٦ .

٤٩- اختبر برنامج الحاسوب، وإذا كان ضروريا أعد كتابته، على الحاسوب المتاح لك.

٥٠- عدل برنامج الحاسوب لإخراج المدرج التكرارى بوحدة مخرجات الرسومات المتاحة لك.

٥١- عدل برنامج الحاسوب لإخراج مدرج تكرارى نسبي، ومدرج تكرارى متجمع صاعد، ومدرج تكرارى متجمع صاعد نسبي بوحدة مخرجات الرسومات المتاحة لك.