أساسيات الإحصاء

FUNDAMENTALS OF STATISTICS

مقدمة INTRODUCTION

تعريف الإحصاء Definition of Statistics

كلمة إحصاء statistics لها معنيان عامان مقبولان :

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

٢ـ 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 فى الفحص يقيس مقاومة الشد لأحد الأجزاء البلاستيكية أو بواسطة باحث تسويق يحدد أفضلية العملاء للألوان. ويمكن أن تجدول بأساليب الورقة والقلم البسيطة أو باستخدام الحاسوب. ويمكن أن يشمل التحليل فحصا مرئيا سريعا أو حسابات مرهقة. والنتائج النهائية تفسر وتقدم للمساعدة فى اتخاذ قرارات خاصة بالجودة.

وهناك مرحلتان للإحصاء :

١ـ إحصاء وصفى أو استدلالى descriptive or deductive ststistics والذى يهدف إلى وصف وتحليل الموضوع أو المجموعة.

٢ـ 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 لا يمكن 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.

تجميع البيانات Collection of Data

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 الرأى العام. ويتم الحصول على البيانات التى تجمع لأغراض مراقبة الجودة بواسطة الملاحظة المباشرة وتقسم إما إلى متغيرات أو خواص. والمتغيرات variables هى صفات الجودة التى تقاس، مثل

الوزن الذى يقاس بالجرامات. والخواص attributes، من ناحية أخرى، هى صفات الجودة التى 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، الذى 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) هى أمثلة لبيانات مستمرة. والمتغيرات التى تعرض فجوات تسمى وثابة descrete . عدد مسامير البرشام غير المطابقة فى مقطورة سفر يمكن أن يكون أى رقم، مثل0 أو3 أو5 أو10 أو96 أو غيرها، إلا أنه لايمكن أن يكون هناك 4.65 مسمار برشام ليس مطابق فى مقطورة معينة. وعموما البيانات المستمرة تكون قابلة للقياس، بينما الوثابة تكون قابلة للعد.

وفى بعض الأحيان يكون مقنعا للبيانات الشفوية أو غير العددية أن تفترض طبيعة المتغير. مثال ذلك، جودة السطح النهائى لقطعة من الاثاث يمكن أن تقسم على أنها ضعيفة، أو متوسطة، أو جيدة. وتقسيم الضعيف، والمتوسط، والجيد يمكن أن يستبدل بقيم عددية مثل1 و2 و3 على التوالى. وبطريقة شبيهة، تشير المنشآت التعليمية إلى رموز المعدلات A، وB، وC، وD، وE، بالقيم العددية، 4 ,3 ,2 ,1 ,0 على التوالى، 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 التى يجب التعبير عنها كخواص. وتكراريا، هذه الصفات التى يحكم عليها بالملاحظة البصرية تقسم على أنها خواص. فالسلك على موتور كهربائى إما أن

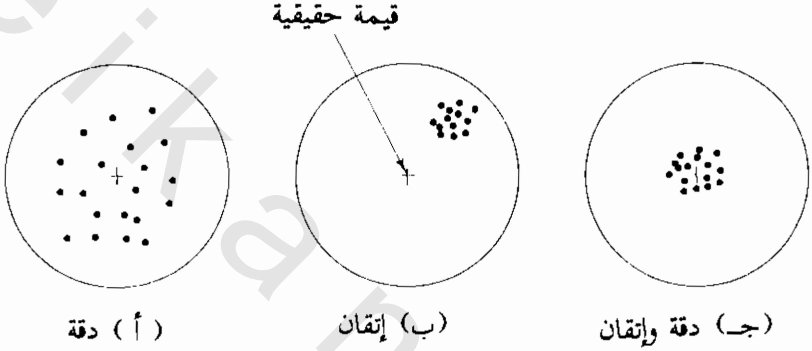
يكون متصلا بنهاية طرفية أو غير متصل، والكلمات فى هذه الصفحة إما أن تكون مكتوبة بهجاء صحيح أو غير صحيح، والمفتاح الترددى إما أن يكون فى وضع التوصيل أو غير التوصيل، والإجابة إما أن تكون صحيحة أو خطأ. والأمثلة المعطاه فى الجملة السابقة تبين مطابقة لمواصفة معينة وعدم مطابقة لنفس المواصفة.

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

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

وقد لا تعطى أجهزة القياس قراءة حقيقية بسبب مشاكل خاصةبالدقة والإتقان. ويبين شكل ٢ ـ ١ ـ أ سلسلة دقيقة من القياسات المتكررة بسبب أن متوسطها يكون قريبا من القيمة الحقيقية، والتى تكون فى المركز. وفى شكل ٢ ـ ١ ـ ب تكون القياسات المتكررة فى السلسلة قريبة جدا من بعضها إلا أنها ليست قريبة من القيمة الحقيقية. ويبين شكل ٢ ـ ١ ـ جـ سلسلة القياسات المتكررة مضغوطة بشدة حول القيمة الحقيقية وهذه القياسات تكون دقيقة ومتقنة.

وتقريب البيانات يتطلب اتباع تحويل معين. ففى 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 الاعداد 0.9535 و0.9536 و0.9537 و0.9538 و0.9539 تكون النتيجة 0.954 ، حيث أن كل الأعداد أقرب إلى0.954 عنها إلى 0.953 . وفى كلمات أخرى، إذا كان الرقم التالى 5 أو أكبر، فيقرب العدد لأعلى.



شكل ٢-١: الفرق بين الدقة والإتقان

عند العمل ببيانات عددية، تكون الأرقام المعنوية significant figures مهمة جدا. والأرقام المعنوية لعدد هى 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 لاتحسب 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

3×10^2 يكون له رقم معنوى واحد، والعدد 3.0×10^2 يكون له رقمان معنويان، والعدد 3.00×10^2 يكون له ثلاثة أرقام معنوية. والأعداد التى تصاحب العد يكون لها عدد لانهائى من الأرقام المعنوية، وحتى العدد 65 يمكن أن يكتب على أنه 65.0000 بأى عدد أصفار، لانهائى، 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.

$$\sqrt{81.9} = 9.05$$

$$6.59 \times 2.3 = 15$$

$32.65 \div 24 = 1.4$ (24 ليس عدد للعد)

$32.65 \div 24 = 1.360$ (24 هو عدد للعد له القيمة 24.00000)

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، لايكون للإجابة النهائية أرقام معنوية أكثر من العدد الذى له أقل عدد للأرقام المعنوية والأمثلة لتوضيح هذه القاعدة هى ما يلى :

$38.26 - 6 = 32$ (6 ليس عدد للعد)

$38.26 - 6 = 32.26$ (6 عدد للعد)

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

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

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

استخدام القواعد السابقة يمنع التعارض فى الإجابات عند أفراد مراقبة الجودة، الا أن بعض الحكم يمكن أن يكون لازما فى بعض الأحيان. وفى أى حالة، يمكن ألا تكون الإجابة أكثر دقة عن البيانات القادمة.

وصف البيانات Describing the Data

فى الصناعة، والأعمال، والمصالح الحكومية كميات البيانات التى جمعت لها أحجام هائلة. حتى عنصر واحد، مثل عدد أجراس الدخان لعدم المطابقة فى35 دفعة بكل منها 1000 ، يمكن أن يمثل مثل هذا الكم الكبير من البيانات الذى يمكن أن يكون مثيرا للخلط أكثر من كونه مفيدا. اعتبر البيانات المبينة فى جدول ٢ ـ ١ . من الواضح، أن هذه البيانات، فى هذه الصورة، تكون صعبة الاستخدام ولا تكون فعالة فى وصف خواص البيانات. وتلزم بعض وسائل تلخيص البيانات لتوضيح أى قيمة أو قيم تتجمع البيانات حولها أو تتشتت أو تبتعد البيانات عنها. وهناك أسلوبان متاحان لتحقيق هذا التلخيص والرسومات والتحليل.

جدول ٢ ـ ١ : عدد أجراس الدخان لعدم المطابقة فى 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 التكرارىfrequency 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 cen tral tendency وقياس التشتت measure of dispersion. وفى بعض الأحيان تستخدم الوسيلتان، وسيلة الرسم والوسيلة التحليلية.

وتوصف هاتان الوسيلتان فى الأقسام التالية من هذا الفصل.

التوزيع التكرارى FREQUENCY DISTRIBUTION

البيانات غير المجمعة Ungrouped Data

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 ـ ١ فى توضيح المفهوم. فالمحلل 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 ـ ٢.

جدول ٢-٢ : عد عدد أجراس الدخان غير المطابقة

التكرار	الجدولة	عدد وحدات عدم المطابقة
13	𝍸 𝍸 |||	0
9	𝍸 ||||	1
5	𝍸	2
4						3
3					4
1			5

الخطوة الأولى هى عمل منظومة array، وهى ترتيب لبيانات عددية خام ترتيبا تصاعديا أو تنازليا طبقا لقيمتها. المنظومة التى لها ترتيب تصاعدى من 0 إلى 5 مبينة فى العمود الأول من جدول ٢ ـ ٢.

الخطوة التالية هى جدولة تكرار كل قيمة عن طريق وضع علامات للعد تحت عمود الجدولة وفى الصف المناسب من جدول ٢ ـ ٢. ابدأ بالأعداد 0 و0 و1 و2 و.. من جدول ٢ ـ ١ واستمر فى وضع علامات العد حتى تتم جدولة البيانات. آخر عمود فى جدول ٢ ـ ٢ هو قيمة عددية لعدد العدادات ويسمى التكرار frequency.

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

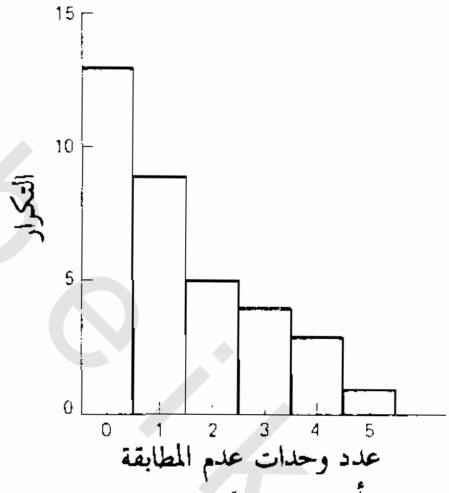
والتوزيع التكرارى هو طريقة مفيدة لرؤية البيانات وهو مفهوم إحصائى أساسى. التفكير فى فئة إعداد يكون لها نوع معين من التوزيع يكون أساسيا فى حل مشاكل مراقبة الجودة. فهناك أنواع مختلفة للتوزيعات التكرارية، ويمكن أن يحدد نوع التوزيع التكرارى طريقة حل المشكلة.

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

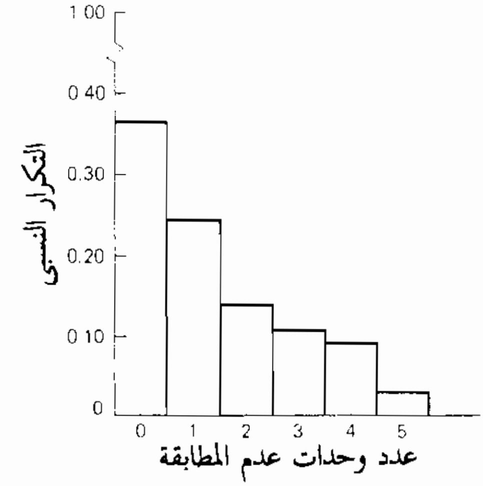
المدرج التكرارى histogram يحتوى على مجموعة من المستطيلات التى تمثل التكرار فى كل فئة. وهو يمثل بالرسم التكرارات للقيم التى تمت ملاحظتها. وشكل ٢ ـ ٢ ـ أ هو مدرج تكرارى لبيانات جدول ٢ ـ ٢. وحيث أن هذا متغير وثاب، فيكون الخط الرأسى فى موقع المستطيل صحيحا نظريا (انظر شكل ٢ ـ ٥). إلا أن المستطيلات شائعة الاستخدام.

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 ـ ٣. والتمثيل بالرسم مبين فى شكل ٢ ـ ٢ ـ 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.

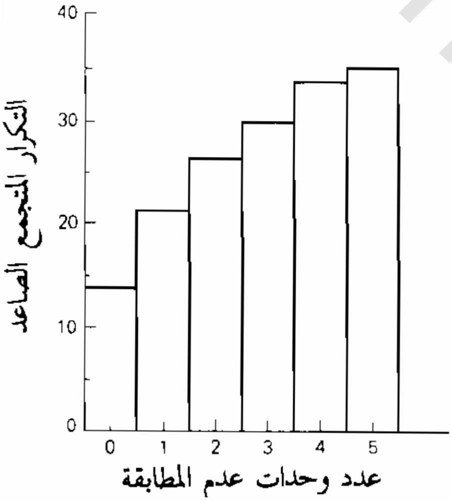
التكرار المتجمع الصاعد يحسب عن طريق جمع تكرار كل قيمة بيانات على مجموع التكرارات لقيم البيانات السابقة. وكما هو مبين فى العمود الرابع من جدول ٢ ـ ٣، التكرار المتجمع الصاعد لوحدات عدم المطابقة التى عددها 0 هو 13، والتى عددها 1 هو $31 = 9 + 22$ والتى عددها 2 هو $27 = 5 + 22$ ، وهكذا. والتكرار المتجمع الصاعد 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 ـ ٢ ـ جـ.



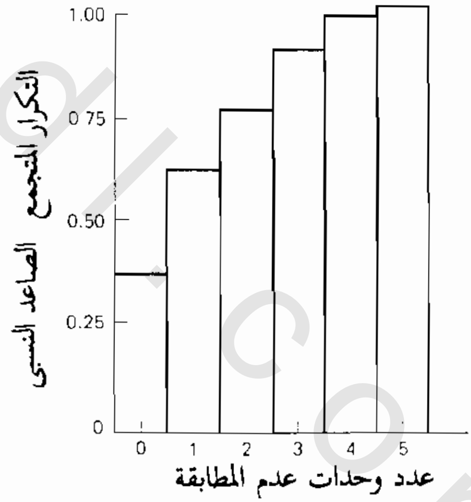
أ ـ مدرج تكرارى



ب ـ مدرج تكرارى نسبى



جـ ـ مدرج تكرارى متجمع صاعد



د ـ مدرج تكرارى لتجمع
صاعد نسبى

شكل ٢٠٢ : تمثيل بالرسومات لبيانات معطاه فى جدولى ٢٠٢ و ٣٠٢

جدول ٢ ـ ٣ : توزيعات تكرارية مختلفة لبيانات معطاه فى جدول ٢ ـ ١

التكرار المتجمع الصاعد النسبى	التكرار المتجمع الصاعد	التكرار النسبى	التكرار	عدد وحدات عدم المطابقة
13 ÷ 35 = 0.37	13	13 ÷ 35 = 0.37	13	0
22 ÷ 35 = 0.63	13 + 9 = 22	9 ÷ 35 = 0.26	9	1
27 ÷ 35 = 0.77	22 + 5 = 27	5 ÷ 35 = 0.14	5	2
31 ÷ 35 = 0.89	27 + 4 = 31	4 ÷ 35 = 0.11	4	3
34 ÷ 35 = 0.97	31 + 3 = 34	3 ÷ 35 = 0.09	3	4
35 ÷ 35 = 1.00	34 + 1 = 35	1 ÷ 35 = 0.03	1	5
		1.00	35	**الإجمالى**

ويحسب التكرار النسبى المتجمع الصاعد عن طريق قسمة التكرار المتجمع الصاعد لكل قيمة بيانات على الإجمالى. هذه الحسابات مبينة فى العمود الخامس من جدول ٢ ـ ٣، وتمثيل الرسومات مبين فى شكل ٢ ـ ٢ ـ د. ويبين الرسم أن تناسب دفعات أجراس الدخان التى 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 التكرارى، إلا أنه لا يقدم معرفة فاحصة بالموضوع.

البيانات المجمعة Grouped Data

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

١ـ إجمع البيانات وقم بإعداد ورقة بالعدcollect data and construct a tally sheet. البيانات المجموعة عن أوزان ١١٠ قضيب صلب مبينة فى جدول ٢ ـ ٤. الخطوة الأولى هى أعداد عدد القيم، كما هو مبين فى جدول ٢ ـ ٥. ولكى نكون أكثر فاعلية، تشفر الأوزان من 2.500 Kg كج، وهذه طريقة مستخدمة فى تبسيط البيانات. لهذا، فالوزن الذى قيمته 31 يكون مكافئا لوزن 2.531 كج (2.500+ 0.031). ويبين تحليل جدول ٢ ـ ٥ أنه تنقل معلومات أكثر إلى المحلل عما تنقل من بيانات جدول ٢ ـ ٤، إلا أن الصورة لاتزال غير واضحة بعض الشىء.

فى هذه المشكلة يوجد 45 فئة، وهى كثيرة جدا ويجب أن تقلل عن طريق عمل التجميع فى خلايا[١]! والخلية هى تجميع داخل حدود محددة للقيم الملاحظة عبر المحور الأفقى للمدرج التكرارى (انظر شكل ٢ ـ ٣). وتجميع البيانات فى خلايا يبسط تمثيل التوزيع، إلا أن بعض التفاصيل تفقد. وعندما يكون عدد الخلايا كبيرا، تنحرف الصورة الحقيقية للتوزيع عن طريق الخلايا التى يكون لها عدد غير كاف من العناصر أو لايكون بها عناصر على الإطلاق. أو، عندما يكون عدد الخلايا صغيرا، يتركز عدد كبير جدا من العناصر فى قلة من الخلايا وينحرف التوزيع أيضا.

عدد الخلايا أو المجموعات فى توزيع تكرارى يكون نوعا من الحكم الشخصى بصورة كبيرة من المحلل. هذا الحكم يكون مبنيا على عدد الملاحظات ويمكن أن يحتاج إلى المحاولة والخطأ لتحديد العدد الأمثل للخلايا. وعموما، يجب أن يقع عدد الخلايا بين 5 و20 . والخطوط الإرشادية العريضة هى كما يلى : 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 و500 ، ويستخدم من 15 إلى 20 خلية عندما يكون عدد الملاحظات أكبر من 500 . ولتوفير المرونة، يتداخل عدد الخلايا

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

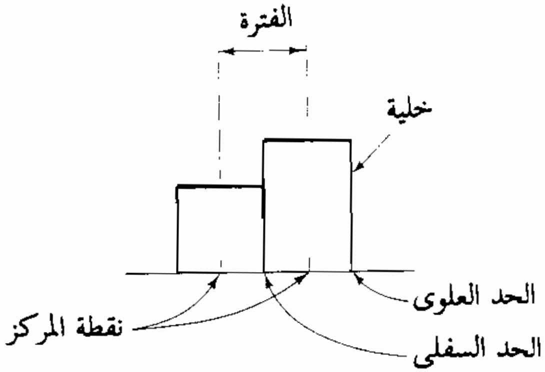
فى الخطوط الإرشادية. وهذا يعنى أن هذه الخطوط الإرشادية ليست متحجرة ويمكن أن تضبط عند الحاجة لتقديم توزيع تكرارى مقبول.

جدول ٢-٤ : 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 ٢-٥ : صفحة أوزان القضبان الصلب (مشفرة من 2.500 Kg)

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



شكل ٢ ـ ٣ : تسمية الخلية

٢ـ حدد المدى 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 = أعلى عدد

X_L = أقل عدد

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

$$R = X_H - X_L$$

$$= 2.575 - 2.531$$

$$= 0.044$$

٣ـ حدد فترة الخلية determine the cell interval . فترة الخلية cell interval هى المسافة بين مركزى خليتين متجاورتين كما هو مبين فى شكل ٢ ـ ٣. وكلما كان ممكنا، 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 فى إيجاد الفترة التى تتفق مع الخطوط الإرشادية.

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

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

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

فترة الخلية التى قيمتها 0.005 مع تسع خلايا تعطى أفضل تقديم للبيانات مبنيا على الخطوط الإرشادية لعدد الخلايا المعطى فى الخطوة الأولى.

طريقة أخرى للحصول على فترة الخلية هى استخدام قاعدة ستورجى Sturgie ، وهى :

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

وبالنسبة إلى مثال المشكلة تكون النتيجة كما يلى:

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

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

٤ـ حدد مراكز الخلايا determine the cell midpoints . أقل مركز خلية يجب أن يحدد ليشمل أقل قيمة بيانات فى خليته. وأبسط طريقة هى اختيار أقل نقطة

بيانات (2.531) كقيمة مركز لأول خلية. ويوصى بهذه الطريقة للقراء المبتدئين لتوهم عملهم فى مجال الإحصاء. وحيث أن الفترة هى 0.005 ، فتوجد خمس قيم بيانات فى كل خلية، لهذا، فإن قيمة المركز 2.533 يمكن أن تستخدم لأول خلية. هذه القيمة يكون لها أقل قيمة بيانات (2.531) فى أول خلية، والتى يكون لها قيم البيانات 2.531 و 2.532 و2.533 و2.534 و2.535 .

اختيار المركز ماهو إلا موضوع يعتمد على الحكم الخاص، وفى 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 فى التوزيع التكرارى. واختيار قيم مراكز مختلفة ينتج عنه توزيعات تكرارية مختلفة ـ وهنا خمسة توزيعات ممكنة. ومراكز الخلايا الثمان المتبقية يتم الحصول عليها عن طريق جمع فترة الخلية على المركز السابق :

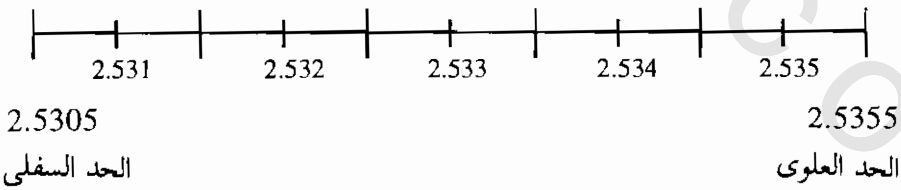
2.538 = 2.533 + 0.005 و 2.543 = 2.538 + 0.005 و= 2.543 + 0.005
2.548 = ... و 2.573 = 2.568 + 0.005 . هذه المراكز مبينة فى جدول ٢ ـ ٦.

جدول ٢ـ٦ : 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

قيمة المركز هى القيمة الأكثر تمثيلا داخل الخلية بافتراض أن عدد الملاحظات فى الخلية يكون كبيرا وأن الفرق فى الحدود لا يكون كبيرا جدا. حتى إذا مالم يتحقق هذا الشرط، فإن عدد الملاحظات الأكبر من والأقل من مركز الخلية يتكرر تساويه. وحتى إذا مالم يكن عدد الملاحظات الأكبر من والأقل من المركز متزنا فى أحد الاتجاهات، فربما يرحل بمقدار عدم الاتزان فى الاتجاه العكسى لخلية أخرى. ويجب أن تكون قيم المراكز بنفس درجة دقة الملاحظات الأصلية.

٥ـ حدد حدود الخلية determine the cell boundaries . حدود الخلية cell boundaries هى قيم الحدود للخلية، ويشار إليها بأنها الحد العلوى والحد السفلى. وكل الملاحظات التى تقع بين الحدين العلوى والسفلى تعد داخل خلية محددة. وتحدد الحدود بحيث لايوجد أى سؤال يخص تحديد موقع إحدى الملاحظات. لهذا، فإن قيم الحدود يكون لها موقع عشرى أو رقم معنوى إضافى فى الدقة وذلك بالنسبة إلى القيم الملاحظة. وحيث أن الفترة تكون فردية، فيوجد عدد متساو من قيم البيانات على كل جانب من جانبى المركز. وبالنسبة إلى الخلية الأولى والتى لها مركز 2.533 وفترة 0.005، توجد قيمتان على كل جانب. لهذا، تحتوى الخلية على القيم 2.531 و2.532 و2.533 و2.534 و2.535 . ولمنع أى فجوات، تتسع الحدود الحقيقية بحوالى منتصف الطريق إلى العدد التالى، ويعطى ذلك القيمتان 2.5305 و2.5355 . ويوضح خط الإعداد التالى هذه القاعدة.



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 مفهوما 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 يلى :

$2.531 + 0.005 = 2.536$ و $2.536 + 0.005 = 2.541$ و ... و $2.566 + 0.005 = 2.571$. ويتم الحصول على الحدود العليا بطريقة شبيهة وهى مبينة فى العمود الأول من جدول ٢ ـ ٦.

٦ـ ضغ تكرار الخلية post the cell frequency . كمية الأعداد فى كل خلية توضع فى عمود التكرار من جدول ٢ ـ ٦. ويبين تحليل جدول ٢ ـ ٥ أنه لأقل خلية يكون هناك : قيمة واحدة من 2.531 واثنتان من2.532 ، وواحدة من2.533 ، 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 ـ ٦. ويعطى التوزيع التكرارى هذا إدراكا أفضل للقيمة المركزية وكيفية تشتت البيانات حول هذه القيمة عن البيانات غير المنظمة أو صفحة العد. ويبين شكل ٢ ـ 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 التكراريين 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 :

١ـ تحديد مقدرة العملية.

٢ـ المقارنة مع المواصفات.

٣ـ اقتراح شكل المجتمع.

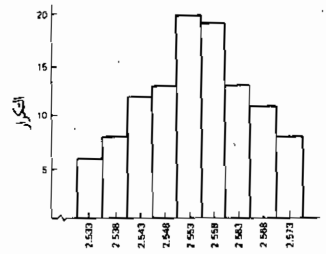
٤ـ تحديد التعارضات فى البيانات مثل الفجوات.

أنواع أخرى لرسومات التوزيع التكرارى

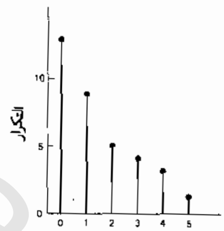
Other Types of Frequency Distribution Graphs

يمكن أن تمثل خريطة الأعمدة التوزيعات التكرارية أيضا، كما هو مبين فى شكل ٢ ـ ٥ ـ أ باستخدام بيانات جدول ٢ ـ 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 ـ ٥ ـ ب مستخدما بيانات جدول ٢ ـ 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.

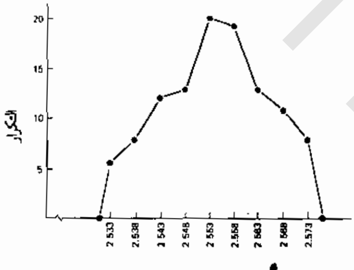


أوزان قضبان الصلب



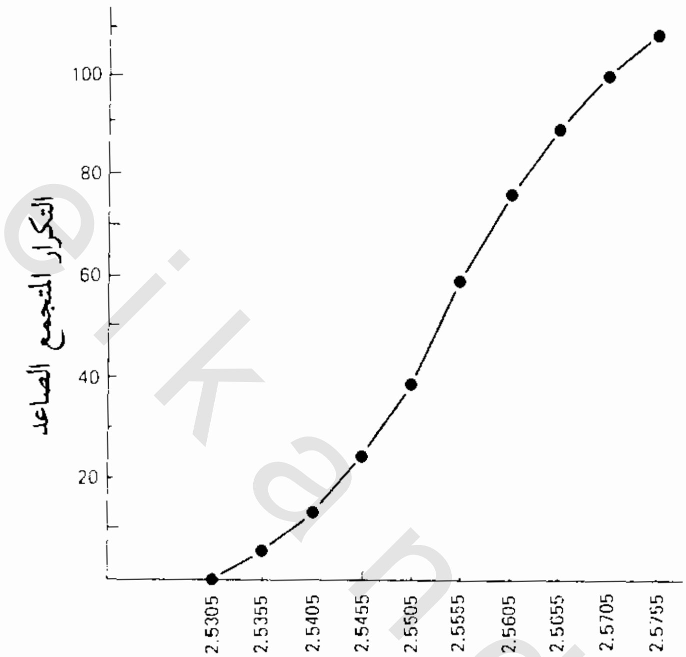
عدد غير المطابق

رسم أعمدة بيانات جدول ٢ ـ ١



أوزان قضبان الصلب
مضلع بيانات جدول ٦٠٢

شكل ٥٠٢ : أنواع أخرى لرسومات التوزيع التكرارى



(جـ) توزيع تكرارى متجمع صاعد لبيانات جدول ٢ ـ 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 ـ ٥ ـ جـ منحنى توزيع تكرارى متجمع صاعد لبيانات جدول ٢ ـ ٦. القيمة المتجمعة الصاعدة لكل خلية ترسم على الرسم وتوصل بخط مستقيم. والحد العلوى الحقيقى للخلية يسمى الإحداثى السينى.

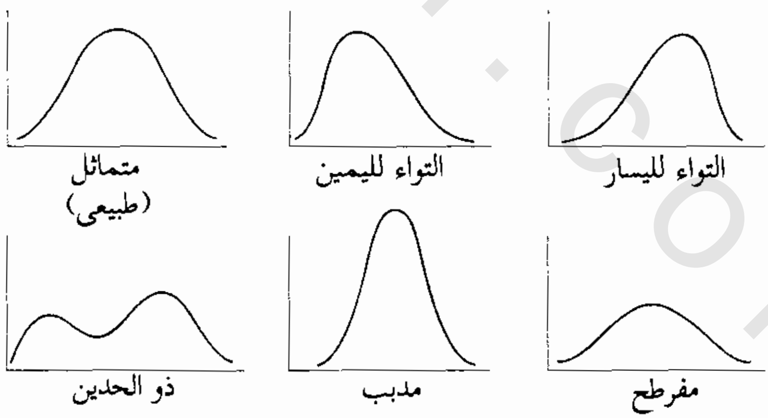
خواص رسومات التوزيع التكرارى

Characteristics of Frequency Distribution Graphs

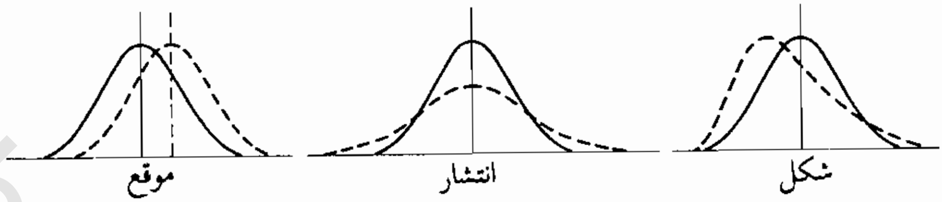
رسومات شكل ٢ ـ ٦ تستخدم منحنيات ملساء بدلا من المستطيلات المصاحبة للمدرج التكرارى. ويمثل المنحنى الأملس توزيعاً تكرارياً للمجتمع، بينما يمثل المدرج التكرارى التوزيع التكرارى للعينة. والفرق بين المجتمع والعينة يناقش فى قسم آخر فى هذا الفصل.

منحنيات التوزيع التكرارى لها خواص معرفة معينة. إحدى هذه الخواص للتوزيع تهم التماثل أو عدم التماثل للبيانات. هل البيانات موزعة توزيعا متساو على كلا الجانبين للقيمة المركزية، أم أنها ملتوية ناحية اليمين أو ناحية اليسار؟ خاصية أخرى تهم عدد المنوالات أو قمم البيانات. ويمكن أن يكون هناك منوالا واحدا، أو اثنين (توزيع ذات الحدين)، أو عدة منوالات. والخاصية الأخيرة خاصة بمدى حدة البيانات peakedness of the data . فعندما يكون المنحنى حادا جدا، فيشار إليه بأنه مدبب leptokurtic ، وعندما يكون مسطحا، 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 ـ ٧.



شكل ٢ـ٦ : خواص التوزيعات التكرارية

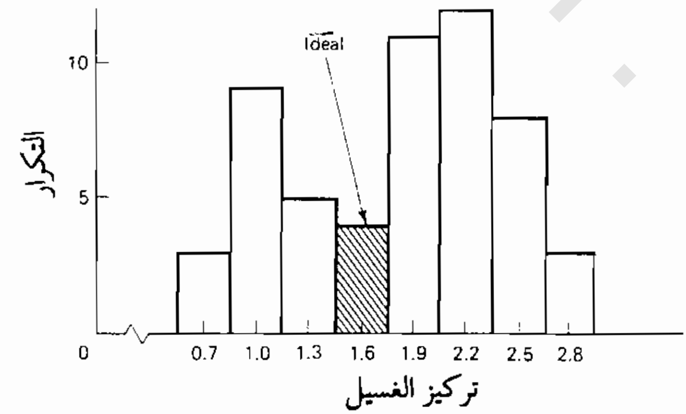


شكل ٢ـ٧ : اختلافات بسبب الموقع، والانتشار، والشكل

تحليل المدرجات التكرارية Analysis of Histograms

يمكن أن يوفر تحليل المدرج التكرارى معلومات خاصة بالمواصفات، وشكل التوزيع التكرارى للمجتمع، ومشكلة مراقبة جودة محددة. ويبين شكل ٢ ـ ٨ مدرجا تكراريا للنسبة المئوية لتركيز الغسيل فى عملية تنظيف أنبوب من الصلب قبل دهانها. ويقع التركيز المثالى بين1.45% و 1.74% ، كما هو مبين بالمستطيل المهشر. والتركيز أقل من 1.45% ينتج جودة ضعيفة، أما التركيز أكبر من 1.75% ، بينما ينتج جودة أكثر من اللازم، يكون مكلفا وبالتالى يقلل الإنتاجية.

ولايلزم إحصاء معقد لتوضيح أن المقاييس التصحيحية تلزم لجعل انتشار التوزيع أقرب إلى القيمة المثالية 1.6% .



شكل ٢ـ٨ : مدرج تكرارى لتركيز الغسيل

تعليقات نهائية Final Comments

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 تكرار إلى الأقل بدلا من الترتيب العددى.

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 ، والتى تناقش فى نهاية الفصل الثالث، تبين الترتيب الذى أنتجت به البيانات ويمكن أن تساعد فى هذا التحليل.

مقاييس النزعة المركزية MEASURES OF CENTERAL TENDENCY

يكون التوزيع التكرارى كافيا للعديد من مشاكل مراقبة الجودة. إلا أنه مع المدى الواسع للمشاكل تصبح طريقة الرسومات غير مرغوب فيها أو أنها تحتاج إلى معلومات إضافية تقدم بواسطة الطرق التحليلية. وتتميز 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 ـ 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 عددية تصف الموقع المركزى للبيانات أو كيف تميل البيانات إلى البناء فى المركز. وهناك ثلاثة مقاييس شائعة الاستخدام : (١) المتوسط، (٢) والوسيط، (٣) والمنوال.

المتوسط Average

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.

١ـ البيانات غير المجمعة ungrouped data . تستخدم هذه الطريقة عندما لاتكون البيانات منظمة. ويمثل المتوسط بالرمز $\overline{X}$، 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

Σ = 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.

مثال لمشكلة EXAMPLE PROBLEM

يختبر أحد القائمين بالفحص قيمة مقاومة خمسة ملفات ويسجل القيم بالاوم (Ω)، وكانت القيم كما يلى :

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

احسب المتوسط

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

$$= \frac{3.35 + 3.37 + 3.28 + 3.34 + 3.30}{5}$$

$$= 3.33\ \Omega$$

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

٢ـ البيانات المجمعة grouped data . عندما تجمع البيانات فى توزيع تكرارى، تسرى الطريقة التالية. صيغة متوسط البيانات المجمعة هى :

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

حيث n = مجموع التكرارات

f_i = تكرار فى 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 التى بكل خلية أكثر من قيمة ملاحظة واحدة كما هو موضح فى مشكلة قضبان الصلب (جدول ٢ ـ ٦). كما أنها تسرى أيضا عندما يكون لكل قيمة ملاحظة، X_i ، تكرارها الخاص بها، f_i، كما هو مبين فى مشكلة أجراس الدخان (جدول ٢ ـ ١). فى هذه الحالة، h تكون عدد القيم الملاحظة.

وفى 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" .

مثال لمشكلة EXAMPLE PROBLEM

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

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

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

$$= 36.1$$ وهى بالألف كيلو متر

لذلك فإن

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

جدول ٢ـ٧: التوزيعات التكرارية لعمر 320 إطار سيارة بالألف كيلو متر

الحدود	نقطة المركز X_i	التكرار f_i	حساب قيمه f_iX_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
الاجمالى		$n = 320$	$\Sigma f_iX_i = 11,549$

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 فى كل خلية تكون موزعة توزيعا غير منتظما فى الخلية. وفى العملية الحقيقية لايكون للاختلاف قيمة كافية للتأثير على دقة المشكلة.

٣ـ المتوسط المرجح weighted average . عندما يدمج عدد من المتوسطات لها تكرارات مختلفة، فيحسب متوسط مرجح weighted average . وصيغة حساب المتوسط المرجح هى كما يلى :

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

حيث: $\bar{x}_w$ = المتوسط المرجح

w_i = ترجيح المتوسط i

EXAMPLE PROBLEM مثال لمشكلة

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

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

$$= \frac{(5)(207) + (6)(203) + (3)(206)}{5 + 6 + 3}$$

$$= 205 \text{ MPa}$$

طريقة المتوسط المرجح هى حالة خاصة لطريقة البيانات المجمعة حيث لا تكون البيانات منظمة فى توزيع تكرارى. وفى 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 مساويا 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 $\bar{x}_i$. وتستخدم نفس المعادلة فى ايجاد :

$\bar{\bar{x}}$ أو $\bar{x}X$ ـ متوسط المتوسطات

$\bar{R}$ ـ متوسط المدى

$\bar{C}$ ـ متوسط العدد غير المطابق

$\bar{S}$ ـ متوسط الانحراف المعيارى للعينة، الخ

الوسيط Median

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

١ـ طريقة البيانات غير المجمعة ungrouped technique. هناك حالتان ممكنتان فى تحديد الوسيط لسلسلة من البيانات غير المجمعة ـ عندما يكون العدد فى السلسلة فرديا، يكون الوسيط مركز هذه القيم. لهذا، فئة الأعداد المرتبة 3 و4 و 5 و6 و8 و8 و10 لها الوسيط 6 وفئة الأعداد المرتبة 22و 24و24 و 24و 30 يكون لها الوسيط. وعندما يكون العدد فى السلسلة زوجيا، يكون الوسيط متوسط العددين الموجودين فى المنتصف. لهذا، فئة الأعداد المرتبة 3 و4 و5 و6 و8 و8 يكون لها الوسيط عبارة عن متوسط 5 و6 أى 5.5 . فإذا كان العددان المتوسطان مثل بعضهما كما فى فئة الأعداد المرتبة 22 و24 و24 و24 و24 و30 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.

٢ـ 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 يلى :

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

حيث : Md = الوسيط

L_m = الحد السفلى للخلية مع الوسيط

n = إجمالى 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 فى جدول ٢ ـ ٧. بالعد من أقل خلية (مركزها 25.0)، يتم الوصول إلى نقطة منتصف الطريق(320/2 = 160) فى 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

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

$$= 35.6 + \left(\frac{\frac{320}{2} - 154}{58} \right) 3$$

$$= 35.9 \qquad \text{(which is in 1000 km)}$$

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

المنوال Mode

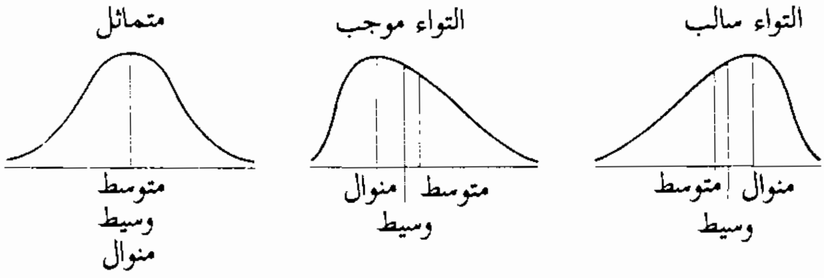
المنوال (MO) mode لفئة أعداد هو القيمة التى تحدث بأكبر تكرار. ومن الممكن للمنوال ألا يوجد فى سلسلة من الأعداد أو أنه يكون له أكثر من قيمة واحدة. للتوضيح سلسلة الأعداد 3 و 3 و 4 و 5 و 5 و 5 و 7 وسلسلة الأعداد 22 و 23 و 25 و 30 و 32 و 36 ليس لها منوال وسلسلة الأعداد 104 و 105 و 105 و 107 و 108 و 109 و 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 فى توزيع تكرارى، فإن مركز الخلية التى 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 لايكون ضروريا، حيث أن المنوال يستخدم أساسا كطريقة فحص لتحديد النزعة المركزية، ولاتلزم دقة أكبر من مركز الخلية.

العلاقة بين مقاييس النزعة المركزية

Relationship Among the Measures of Central Tendency

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



شكل ٢-٩: العلاقة بين المتوسط والوسيط، والمنوال

المتوسط هو المقياس الأكثر استخداما للنزعة المركزية. ويستخدم عندما يكون التوزيع متماثلا أو غير مميز انحرافه لليمين أو لليسار، وعندما يراد حساب الإحصاء الإضافى مثل قياسات التشتت، وخرائط المراقبة، وغيرها اعتمادا على المتوسط، وعندما تلزم قيمة مستقرة للإحصاء الاستدلالى.

ويصبح الوسيط مقياسا فعالا للنزعة المركزية عندما يكون التوزيع ملتويا التواء 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، فى سلسلة أعداد مثل 12و 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 وتقريبى للنزعة المركزية. لهذا، يوجد منوال المدرج التكرارى بسهولة عن طريق الفحص البصرى. بالإضافة إلى ذلك، يستخدم المنوال فى وصف القيم التقليدية أكثر للتوزيع، مثل منوال عمر مجموعة معينة.

وهناك مقاييس أخرى للنزعة المركزية وهى الوسط الهندسى، والوسط التوافقى، والوسط التربيعى. ولا تستخدم هذه المقاييس فى مراقبة الجودة.

مقاييس التشتت MEASURES OF DISPERSION

مقدمة Introduction

لقد نوقشت فى القسم السابق طرق وصف النزعة المركزية للبيانات. ووسيلة أخرى للإحصاء تتكون من مقاييس التشتت measures of dispersion ، والتى 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 فى أقسام الطلاء المعدنى والتجميع لأحد المصانع يكون لهم متوسط أجر أسبوعى متطابق قيمته 225.36 دولار، إلا أن قسم الطلاء المعدنى 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 المعدنى.

ومقاييس التشتت المناقشة فى هذا القسم هى المدى، والانحراف المعيارى، والتباين. والمقاييس الأخرى مثل وسط التباين الحسابى والانحراف الربيعى لاتستخدم فى مراقبة الجودة.

المدى Range

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

$$R = X_H - X_L$$

حيث : R = المدى

X_H = أعلى ملاحظة فى السلسلة

X_L = أقل ملاحظة فى السلسلة

مثال لمشكلة EXAMPLE PROBLEM

إذا كان أعلى أجر أسبوعى فى قسم التجميع هو 280.79 دولار وأقل أجر أسبوعى هو 173.54 دولار، حدد المدى.

$$R = X_H - X_L$$
$$= \$280.79 - \$173.54$$
$$= \$107.25$$

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

الانحراف المعيارى Standard Diviation

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

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

حيث : S = الانحراف المعيارى للعينة

X_i = القيمة الملاحظة

$\bar{x}$ = المتوسط

n = عدد القيم الملاحظة

يستخدم جدول ٢ ـ ٨ فى توضيح مفهوم الانحراف المعيارى. العمود الأول (X_i) يعطى ست قيم ملاحظة بالكيلوجرام، ومن هذه القيم يتم الحصول على المتوسط $\bar{x} = 3.0$ العمود الثانى $(X_i - \bar{x})$ ، هو انحراف القيم الملاحظة الفردية عن المتوسط. إذا ما جمعنا الانحرافات، تكون النتيجة صفرا، وهذا هو الحال دائما، إلا أن هذا لا يقود إلى قياس التشتت. إلا أنه إذا تم تربيع الانحرافات، فإنها تصبح موجبة ويصبح مجموعها أكبر من صفر. والحسابات مبينة فى العمود الثالث، $(X_i - \bar{x})^2$ ، مع المجموع الناتج مساويا 0.08 ، والذى يتغير طبقا للقيم الملاحظة. ومتوسط مربعات الانحرافات يمكن أن يوجد عن طريق القسمة على n ، إلا أنه لأسباب نظرية فإننا نقسم على $n-1$[١]. لهذا،

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

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

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

جدول ٨ـ٢: تحليل الانحراف المعيارى

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 التربيعى، فتصبح النتيجة بنفس وحدات القيم الملاحظة. وتكون الحسابات كما يلى :

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

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 ـ طريقة البيانات غير المجمعة ungrouped technique . الصيغة المستخدمة فى تعريف الانحراف المعيارى يمكن أن تستخدم للبيانات غير المجمعة. إلا أنه هناك صيغة أكثر راحة تستخدم فى أغراض الحسابات :

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

مثال لمشكلة EXAMPLE PROBLEM

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

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

$$= \sqrt{\frac{6\,(231.26) - (37.2)^2}{6\,(6-1)}}$$

$$= 0.35\%$$

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

٢ ـ طريقة البيانات المجمعة grouped data . عندما تكون البيانات مجمعة فى توزيع تكرارى، تطبق الطريقة التالية. صيغة الانحراف المعيارى للبيانات المجمعة هى كما يلى :

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

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 "fX2" ، كما هو مبين فى جدول ٢ ـ ٩. ويجب تذكر أن عمود "fX" يلزم لحسابات المتوسط، لهذا فإنه يلزم عمود إضافى وحيد لحساب الانحراف المعيارى. والطريقة مبينة بمثال المشكلة التالى :

جدول ٢ ـ ٩: سرعات سيارات الركاب (كم/ساعة) أثناء فترة من 15 دقيقة.

الحدود	نقطة المركز X_i	التكرار f_i	الحسابات f_iX_i	$f_iX_i^2$
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 ٢ • ٩ لسرعات سيارات الركاب أثناء فترة من 15 دقيقة، حدد المتوسط والانحراف المعيارى.

$$\bar{x} = \frac{\sum_{i=1}^{h} f_i X_i}{n} = \frac{9354}{96} = 97.4 \text{ km/h}$$

$$S = \sqrt{\frac{n \sum_{i=1}^{h} (f_i X_i^2) - \left(\sum_{i=1}^{h} f_i X_i\right)^2}{n(n-1)}} = \sqrt{\frac{96(920,742) - (9354)^2}{96(96-1)}} = 9.9 \text{ km/h}$$

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 المعيارى لعينة القيم الملاحظة. وتستخدم نفس الصيغة فى إيجاد :

$S_{\bar{x}}$ – الانحراف المعيارى لمتوسطات العينات

S_p – الانحراف المعيارى لنسب العينات

S_R – الانحراف المعيارى لمجموعات مدى العينات

S_S – الانحراف المعيارى للانحرافات المعيارية للعينات، إلخ.

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

العلاقة بين مقاييس التشتت

Relationship Between the Measures of Dispersion

فى مراقبة الجودة يكون المدى مقياسا شائعا جدا للتشتت، ويستخدم فى أحد خرائط المراقبة الرئيسية. وتكمن الميزة الأولية للمدى فى تقديمه معرفة بإجمالى 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 الاسهل حدوث قراءات كبيرة جدا وأخرى صغيرة جدا. ويقترح أن يكون استخدام المدى محددا بعشر ملاحظات كحد أقصى.

ويستخدم الانحراف المعيارى عندما يكون مطلوبا مقياسا أكثر إتقانا. كما أنه المقياس الأكثر شيوعا للتشتت فى عمل مراقبة الجودة ويستخدم عندما يراد عمل حسابات إحصائية تالية. عندما يكون للبيانات قيمة مرتفعة ارتفاعا غير عادى أو منخفضة انخفاضا غير عادى، فيكون الانحراف المعيارى مرغوبا أكثر عن المدى.

مقاييس أخرى OTHER MEASURES

هناك مقياسان آخران يتكرر استخدامهما فى تحليل مجموعة من البيانات ـ الالتواء والتفرطح.

الالتواء Skewness

كما سبق ذكره، الالتواء skewness هو فقدان التماثلية للبيانات. والصيغة معطاه كما يلى[١]:

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

حيث a_3 تمثل الالتواء.

والالتواء عبارة عن عدد يذكر حجمه مدى الترحيل من التماثلية. إذا كانت قيمة a_3 هى صفر، تكون البيانات متماثلة، وإذا كانت أكبر من صفر (موجبة)، تكون البيانات ملتوية لليمين، وهذا يعنى وجود ذيل طويل ناحية اليمين، وإذا كانت أقل من صفر (سالبة)، فتكون البيانات ملتوية ناحية اليسار، وهذا يعنى وجود ذيل طويل ناحية اليسار. انظر شكل ٢ ـ ٩ لرؤية تمثيل بيانى للالتواء. وتشمل قيم 1+ أو 1- توزيعا متماثلا تماثلاقويا.

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

EXAMPLE PROBLEM مثال لمشكلة

حدد التواء التوزيع التكرارى لجدول ٢ ـ ١٠. حُسِبَ المتوسط وكذلك الانحراف المعيارى وكانت قيمتهما 7.0 و2.32 ، على التوالى .

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

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

$$= -0.42$$

جدول ٢ـ١٠ : 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 البصرى لعمودى x و f 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 المنوال. وتقدم قيمة الانحراف معلومات خاصة بشكل توزيع المجتمع. مثال ذلك، يكون للتوزيع الطبيعى التواء صفر.

التفرطح Kurtosis

كما سبق ذكره من قبل التفرطح هو حالة قمة البيانات. وتعطى الصيغة على النحو التالى(١):

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

حيث a_4 تمثل التفرطح

التفرطح هو قيمة بدون وحدات وتستخدم كمقياس مقارن لارتفاع القمة فى توزيعين. انظر شكل ٢ ـ ٦ وشكل ٢ ـ ١١ لترى التمثيل البيانى للتفرطح.

مثال لمشكلة EXAMPLE PROBLEM

حدد التفرطح للتوزيع التكرارى لجدول ٢ ـ ١٠، والذى له $\bar{X} = 7.0$ و $S = 2.32$.

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

$$= \frac{2430 / 31}{2.32^4}$$

$$= 2.70$$

قيمة التفرطح لا توفر أى معلومات فى حد ذاتها ـ فيجب أن تقارن مع توزيع آخر. واستخدام قيمة التفرطح هو نفسه مثل الانحراف ـ حجم عينة n كبير، وتوزيع أحادى المنوال. مثال ذلك، التوزيع الطبيعى، توزيع متوسط التفرطح mesokurtic، له

(١) هذه الصيغة هى 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 الطبيعى، توزيع مدبب leptokurtic ، وإذا كانت $a < 3$ فيكون ارتفاع التوزيع أقل فى قمته عن التوزيع الطبيعى، 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

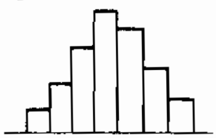
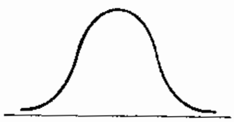
عند هذه النقطة، يكون من المرغوب فيه فحص مفهوم المجتمع والعينة. ولكى نعد توزيعا تكراريا لأوزان قضبان من الصلب، يتم اختيار جزء صغير، أو عينة sample، لتمثيل كل قضبان الصلب. وبالمثل، البيانات المجمعة الخاصة بسرعات سيارات الركاب تمثل جزءا صغيرا فقط لكل سيارات الركاب. والمجتمع هو التجميع الكلى للقياسات، وفى 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 لانهائيا أو لانهائيا 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 فى حالة الأدوية الجديدة. وفعليا، فمحلل المجتمع كله لن يكون أكثر دقة من المعاينة. فقد تبين أن فحص يدوى100% ليس بنفس دقة المعاينة. وربما يرجع ذلك إلى الحقيقة بأن الضجر والتعب يتسببان فى حكم مسبق للعنصر المفحوص بأنه مقبول من قبيل الفاحص.

وعند تحديد مجتمع، يستخدم الحرف اليونانى المناظر. لهذا، يكون لمتوسط العينة الرمز $\bar{x}$ ، ويكون للوسط الحسابى للمجتمع الرمز μ (ميو). لاحظ أن كلمة «متوسط» تغيرت إلى «وسط حسابى» عند استخدام المجتمع. والرمز $\bar{x}o$ هو القيمة النمطية أو الدليل النمطى. وتبنى المفاهيم الرياضية على μ، وهى القيمة الحقيقية $\bar{x}o$ تمثل مكافئ عملى بغرض استخدام المفاهيم. والانحراف المعيارى للعينة له الرمز s ، وللانحراف المعيارى للمجتمع الرمز σ (سيجما). والرمز s هو قيمة نمطية أو إشارة نمطية وله نفس العلاقة مع σ مثل علاقة $\bar{x}o$ مع μ. وقد لا تعرف قيمة المجتمع الحقيقية على الإطلاق، لهذا، أحيانا يستخدم الرمزان $\hat{\mu}$ و $\hat{\sigma}$ للدلالة على «تقدير لـ». وتوجد مقارنة بين العينة والمجتمع فى جدول ٢ ـ ١١. وسوف تذكر مقارنات إضافية عند حدوثها فيما بعد.

جدول ٢ ـ ١١: مقارنة بين العينة والمجتمع

العينة	المجتمع
إحصائية	معلمة
$\bar{x}$ – المتوسط	$\bar{x}o$ μ الوسط الحسابى
s – الانحراف المعيارى للعينة	(s_o) σ الانحراف المعيارى
	

الهدف الأساسى من اختيار العينة هو تعلم شيء 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.

ويبين جدول ٢ ـ ١٢ نتائج تجربة توضح العلاقة بين العينات والمجتمع. ففى الحاوى الذى به 800 كرة زرقاء، و200 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 ومايجب 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 ٢ ـ ١٢: 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

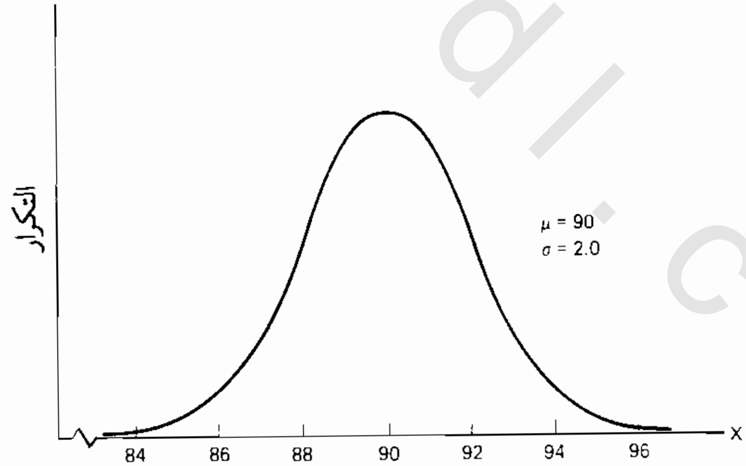
بينما حدثت استدلالات عن المجتمع من العينات، فإنه من الحقيقى أن معرفة المجتمع تقدم معلومات لتحليل العينة. لهذا، يكون ممكنا تحديد ما إذا كانت العينة قد سحبت من مجتمع معين. هذا المفهوم يكون ضروريا فى فهم نظرية خرائط المراقبة. وقد تأجلت مناقشة تفصيلية أكثر إلى الفصل الثالث.

المنحنى الطبيعى THE NORMAL CURVE

الوصف Description

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 الطبيعى normal curve أو توزيع جاوس Gaussian distribution . والمنحنى الطبيعى هو توزيع متماثل، يشبه الجرس، له منوال واحد، وله نفس القيمة لكل من الوسط الحسابى، والوسيط، والمنوال.

المنحنى الطبيعى أو التوزيع الطبيعى يعد من المدرج التكرارى. وكلما كبر حجم عينة المدرج التكرارى وازداد كبرا، 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 ـ ١٠.



شكل ٢ ـ ١٠ : 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 الظباء، وأطوال الأفراد الآدميين تتبع منحنىً 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 منحنىً طبيعيا. عند اعتبار أطوال أفراد آدميين، يمكننا أن نتوقع أن تكون نسبة مئوية صغيرة منهم لها أطوال كبيرة جدا ونسبة صغيرة منهم لها أطوال قصيرة جدا، مع تقارب معظم الأطوال حول المتوسط. والمنحنى الطبيعى هو وصف جيد للتغيرات التى تحدث لمعظم خواص الجودة فى الصناعة التى تكون أساسا للعديد من أساليب مراقبة الجودة.

كل التوزيعات الطبيعية للمتغيرات المستمرة يمكن أن تحول إلى توزيع طبيعى نمطى (انظر شكل ٢ ـ ١٠) عن طريق استخدام قيمة طبيعية نمطية Standard-ized normal value. مثال ذلك، اعتبر القيمة 92 أوم فى شكل ٢ ـ 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 ـ ١١.

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

والصيغة للمنحنى الطبيعى النمطى هى :

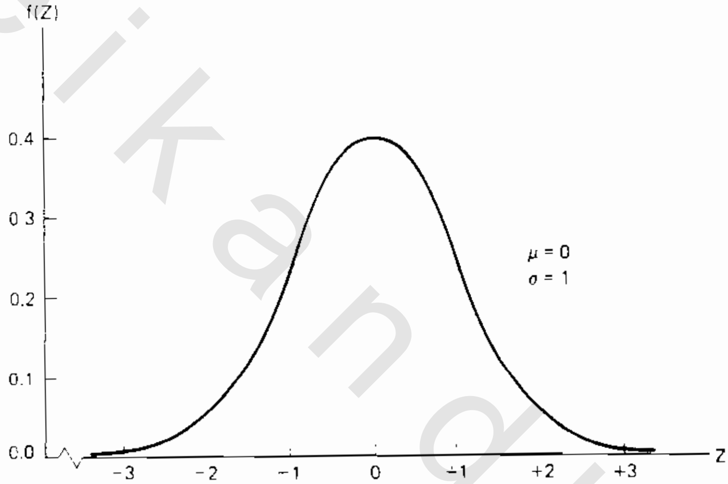
حيث : $\pi = 3.142593$

$$e = 2.71828$$

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

يبين شكل ٢ ـ ١١ المنحنى النمطى بوسط حسابى صفر وانحراف معيارى 1. ويلاحظ أن المنحنى مقارب عند 3- = Z و Z = 3 .

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



شكل ٢ ـ ١١ : توزيع طبيعى نمطى بوسط حسابى صفر وانحراف معيارى1 .

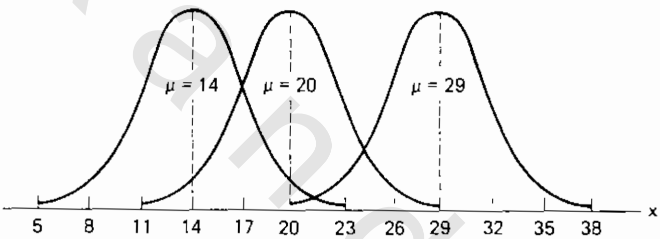
ويمكن الإشارة إلى التوزيع الطبيعى بأنه توزيع احتمالى طبيعى. وبينما يكون أكثر توزيعات المجتمعات أهمية، إلا أنه هناك عدد من التوزيعات الأخرى للمتغيرات المستمرة. فهناك أيضا عدد من التوزيعات الاحتمالية للمتغيرات الوثابة. وهذه التوزيعات تناقش فى الفصل الرابع.

العلاقة بالوسط الحسابى والانحراف المعيارى

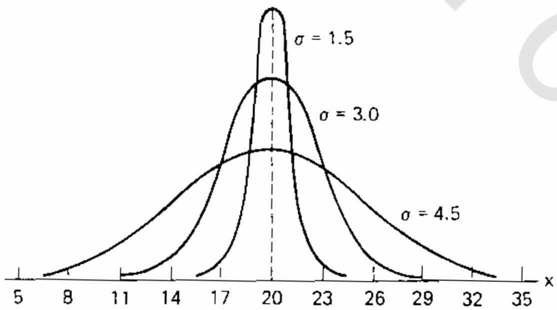
Relationship to the Mean and Standard Diviation

كما هو مبين بالصيغة الخاصة بالمنحنى الطبيعى النمطى، توجد علاقة محددة بين الوسط الحسابى، والانحراف المعيارى، والمنحنى الطبيعى. يبين شكل ٢ ـ ١٢

ثلاثة منحنيات طبيعية لها قيم وسط حسابى مختلفة، ويلاحظ أن التغير الوحيد يوجد فى الموقع فقط. ويبين شكل ٢ ـ ١٣ ثلاثة منحنيات طبيعية لها نفس الوسط الحسابى ولكنها لها انحرافات معيارية مختلفة. ويوضح الشكل القاعدة بأنه كلما كبر الانحراف المعيارى، كلما تسطح المنحنى (تكون البيانات مشتتة أكثر)، وكلما قل الانحراف المعيارى، كلما ازدادت قمة المنحنى (يضيق تشتت البيانات). وإذا كان الإنحراف المعيارى صفرا، تكون كل القيم متطابقة مع الوسط الحسابى ولا يكون هناك منحنى على الإطلاق.



شكل ٢ ـ ١٢: منحنى طبيعى بأوساط حسابية مختلفة وانحرافات معيارية متطابقة.

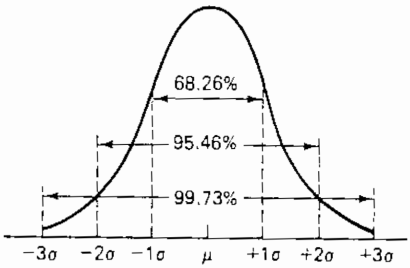


شكل ٢ ـ ١٣: منحنى طبيعى بانحرافات معيارية مختلفة لكن بأوساط حسابية متطابقة.

يعرف المنحنى الطبيعى تعريفا كاملا بواسطة الوسط الحسابى والانحراف المعيارى للمجتمع. وكذلك، كما هو مبين فى شكلى ٢ ـ ١٢ و ٢ ـ ١٣، هاتان المعلمتان تكونان مستقلتين. وفى كلمات أخرى، لا يؤثر التغير فى أحدهما على الآخر.

وتوجد علاقة بين الانحراف المعيارى والمساحة تحت المنحنى كما هو مبين فى شكل ٢ ـ ١٤. ويبين الشكل أنه فى المنحنى الطبيعى 68.26% من العناصر تكون مشمولة بين حدى إنحراف معيارى 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 المالانهاية (الحد السالب والحد الموجب). وهذه النسب تكون حقيقية بغض النظر عن شكل المنحنى الطبيعى. والحقيقة أن 99.73% من العناصر تكون مشمولة بين حدى ثلاثة انحرافات معيارية هى أساس خرائط المراقبة التى تناقش فى الفصل الثالث.



شكل ٢ ـ ١٤: النسب المئوية للعناصر المشمولة بين قيمتين معينتين للانحراف المعيارى

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 الطبيعى» هو جدول قراءة من ناحية اليسار،[1] وهذا يعنى ان المساحات المعطاة تكون لجزء المنحنى من ناقص مالانهاية 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

μ = الوسط الحسابى

σ = الانحراف المعيارى للمجتمع

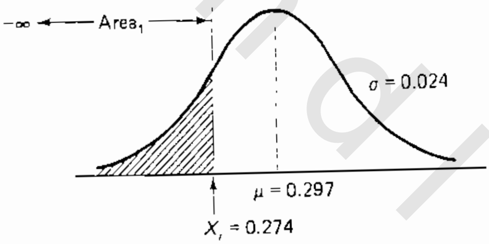
بعد ذلك، باستخدام قيمة Z المحسوبة تتحدد المساحة على يسار X_i من جدول أ. لهذا، إذا كانت Z المحسوبة 1.76-، تكون قيمة المساحة 0.0392 . وحيث أن المساحة الكلية تحت المنحنى هى 1.000 ، فإن قيمة 0.0392 من المساحة يمكن أن تتغير إلى نسبة مئوية للعناصر تحت المنحنى عن طريق نقل العلامة العشرية موقعين إلى اليمين.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

(١) فى بعض الكتب يرتب جدول المساحات تحت المنحنى الطبيعى 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 .

مثال لمشكلة EXAMPLE PROBLEM

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



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

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

$$= -0.96$$

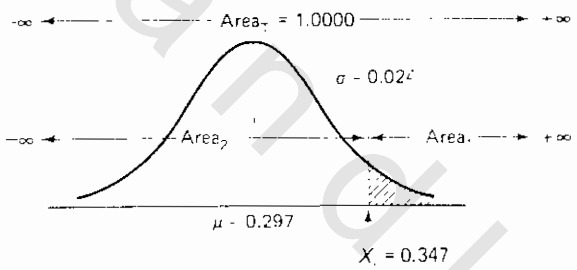
من جدول أ توجد قيمة Z وهى 0.96- ، والمساحة = 0.1685 أو 16.86% .

لهذا، 16.85% من البيانات تكون أقل من 0.274

مثال لمشكلة EXAMPLE PROBLEM

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

حيث أن جدول أ هو جدول قراءة من اليسار، فيتطلب 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.



$$Z_2 = \frac{X_i - \mu}{\sigma}$$

$$= \frac{0.347 - 0.297}{0.024}$$

$$= + 2.08$$

من جدول أ يوجد أن لقيمة Z المساوية 2.08 +،

$$Area_2 = 0.9812$$

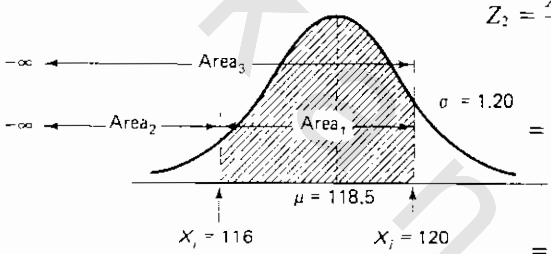
$$Area_1 = Area_T - Area_2 = 1.000 - 0.9812 = 0.018$$ أو 1.88%

لهذا، 1.88% من البيانات تقع بعد 0.347 كج.

مثال لمشكلة EXAMPLE PROBLEM

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

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



$$Z_2 = \frac{X_i - \mu}{\sigma} = \frac{116 - 118.5}{1.20} = -2.08$$

$$Z_3 = \frac{X_i - \mu}{\sigma} = \frac{120 - 118.5}{1.20} = +1.25$$

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

$$Area_1 = Area_3 - Area_2$$

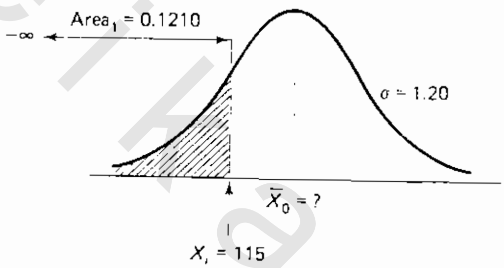
$0.8944 - 0.0188 = 0.8756$ أو 87.56%

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

مثال لمشكلة EXAMPLE PROBLEM

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

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



$$Z_1 = \frac{X_i - \bar{X}_0}{\sigma}$$

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

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

لهذا، يجب أن يتمركز الوسط الحسابي 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 مبنيا على قيم μ و σ ، إلا أن $\bar{x}_0$ و s_0، يمكن أن تعوض باشتراط أنه هناك إشارات بأن التوزيع طبيعي. ويوضح آخر مثال لمشكلة استقلالية μ و σ. ولا يؤثر التغيير البسيط في مركزة العملية على التشتت.

اختبارات الطبيعية TESTS OF NORMALITY

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 كاى تربيع تطبق أيضا مع بعض التعديل على توزيعات المجتمع الأخرى.

المدرج التكرارى Histogram

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 ـ 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.

الانحراف والتفرطح Skewness and Kurtosis

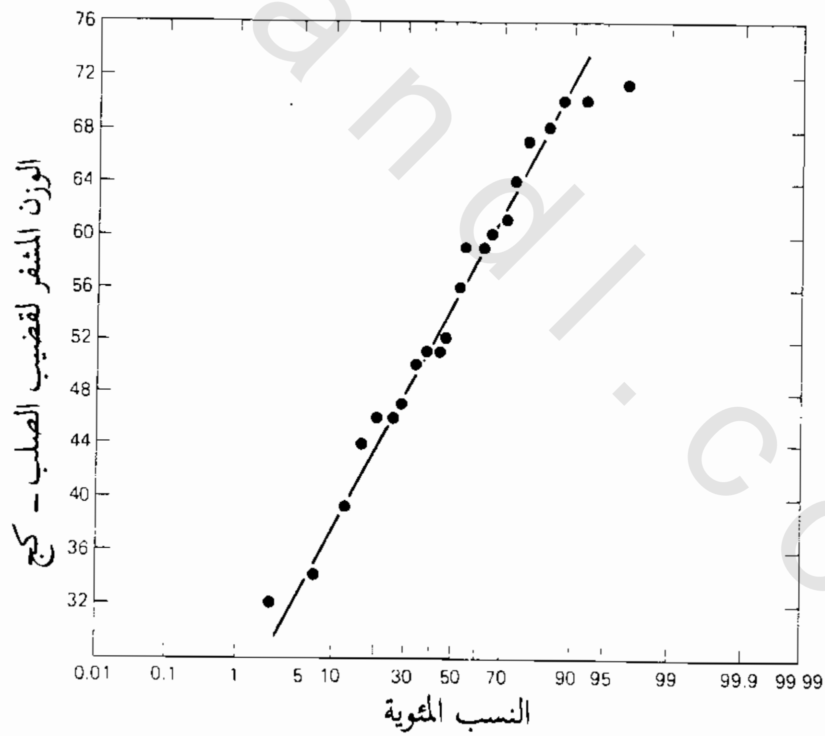
مقاييس الانحراف والتفرطح هى اختبار آخر للطبيعية. من بيانات قضبان الصلب الموجودة فى جدول ٢ ـ ٦، نجد أن $a_3 = -0.11$ وأن $a_4 = 2.19$. تحدد هذه القيم أن البيانات منحرفة قليلا ناحية اليسار إلا أنها قريبة من القيمة الطبيعية صفر، كما أن البيانات ليست لها قمة مثل التوزيع الطبيعى، والتى لها قيمة $a_4 = 3.0$.

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.

رسومات الاحتمالات Probability Plots

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

١ـ رتب البيانات order the data . البيانات من العمود الأول من جدول ٢ـ ٤ تستخدم فى توضيح المفهوم. تسجل كل ملاحظة كما هو مبين فى جدول ٢ـ١٣ من الأقل وحتى الأكبر. والملاحظات المزدوجة تسجل كما هو مبين بواسطة القيمة 46.



شكل ٢ـ١٥ : رسومات الاحتمالات للبيانات الموجودة فى جدول ٢ ـ ١٣

جدول ٢ـ١٣ : بيانات عن أوزان قضبان الصلب لرسومات الاحتمالات

الملاحظة X_i	الرتبة i	موقع الرسم	الملاحظة X_i	الرتبة 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

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

٣ـ 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 = 22/ (0.5 - 1) 100 ، أى 2.3% ، وتحسب المواقع الأخرى بنفس الطريقة وتوضع فى جدول ٢ ـ ١٣ .

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

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

٦ـ حاول أن تضبط بالعين «أفضل» خط attempt to fit by eye a "best" 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.

٧ـ 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 تقريبية55. 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.

كاى تربيع Chi-Square

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

ومن المهم للمحلل أن يفهم أن أى طريقة من هذه الطرق لاتثبت أن البيانات موزعة توزيعا طبيعيا. وإنما يمكننا أن نستخلص فقط أنه لايوجد دلائل على أن البيانات لايمكن أن تعامل كما لو كانت موزعة توزيعا طبيعيا.

برنامج حاسوب COMPUTER PROGRAM

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 ـ 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 ـ 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(I) < 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

مشاكل PROBLEM

١ــ قرب الأعداد التالية إلى أقرب موقعين عشريين :

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

٢ـ نفذ العملية المحددة مع إعطاء الإجابة بالأعداد الصحيحة للأرقام المعنوية.

أ ـ (34.6) (8.20) ب ـ (0.035) (635) ج ـ 3.8735/6.1

د ـ 5.362/6 (6 هو عدد للعد).

٣ـ نفذ العملية المحددة مع إعطاء الإجابة بالأعداد الصحيحة للأرقام المعنوية.

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

٤ـ لقد سجل أحد لاعبى كرة السلة فى 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 ـ اعمل استمارة (ورقة) عد فى ترتيب تصاعدى

ب ـ مستخدما البيانات السابقة، إرسم مدرجا تكراريا.

٥ـ تملأ إحدى الشركات زجاجات من الشامبو وتحاول أن تحتفظ بوزن محدد لمنتجها. ويقدم الجدول أوزان 110 زجاجة من زجاجات الشامبو التى 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

٦ـ فيما يلى 125 قراءة تم الحصول عليها من إحدى المستشفيات بواسطة محلل لإحدى دراسات الوقت والحركة الذى أخذ خمس قراءات كل يوم لمدة25 يوما. اعمل ورقة للعد. ارسم جدولا مبينا مراكز الخلايا، والتكرارات الملاحظة. ارسم مدرجا تكراريا.

اليوم	استمرارية وقت العملية 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

٧ـ اختبرت القوة النسبية لعدد 150 سبيكة لحام فضية، والنتائج معطاة فى الجدول التالى. عد هذه الأرقام ورتبها فى توزيع تكرارى. حدد فترة الخلية والعدد التقريبى للخلايا. ارسم جدولا مبينا 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 ـ باستخدام بيانات المشكلة رقم ٤، ارسم :

أ ـ مدرجا تكراريا نسبيا.

ب ـ مدرجا تكراريا متجمعا صاعدا.

ج ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

٩ـ باستخدام بيانات المشكلة رقم ٥، ارسم :

أ ـ مدرجا تكراريا نسبيا.

ب ـ مدرجا تكراريا متجمعا صاعدا.

ج ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

١٠ـ باستخدام بيانات المشكلة رقم ٦ ، ارسم :

أ ـ مدرجا تكراريا نسبيا.

ب ـ مدرجا تكراريا متجمعا صاعدا.

ج ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

١١ـ باستخدام بيانات المشكلة رقم ٧، ارسم :

أ ـ مدرجا تكراريا نسبيا.

ب ـ مدرجا تكراريا متجمعا صاعدا.

ج ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

١٢ـ ارسم خريطة قضبان للبيانات الموجودة في :

أ ـ المشكلة رقم ٤.

ب ـ المشكلة رقم ٥.

١٣ـ باستخدام بيانات المشكلة رقم ٦ ، ارسم :

أ ـ مضلعا تكراريا.

ب ـ منحنى متجمع صاعد.

١٤ـ باستخدام بيانات المشكلة رقم ٧، ارسم :

أ ـ مضلعا تكراريا.

ب ـ منحنى متجمع صاعد.

١٥ـ يختبر أحد التقنيين الكهربائيين شدة التيار الداخل لأحد المنازل السكنية وحصل على الخمس قراءات التالية : 115 و 113 و 121 و 115 و 116. ما المتوسط؟

١٦ـ قام أحد العاملين بثمان رحلات لتحصيل مقطورة. إذا كانت مسافات الرحلات بالمتر وكانت كما يلى : 25.6 و 24.8 و 22.6 و 21.3 و 19.6 و 18.5و 16.2 و 15.5 ، ما المتوسط ؟

١٧ـ 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

١٨ـ 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 ..

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

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

٢١ـ حدد الوسيط للأعداد التالية :

أ ـ 22 و 11و 15 و 8 و 18.

ب ـ 35 و 28 و 33 و 38 و 43 و 36 .

٢٢ـ حدد الوسيط لكل مما يلى :

أ ـ التوزيع التكرارى للمشكلة رقم ١٧ .

ب ـ التوزيع التكرارى للمشكلة رقم ١٨ .

ج ـ التوزيع التكرارى للمشكلة رقم ٢٨ .

د ـ التوزيع التكرارى للمشكلة رقم ٣٠ .

هـ ـ التوزيع التكرارى للمشكلة رقم ٦ .

و ـ التوزيع التكرارى للمشكلة رقم ٧.

٢٣ـ بمعرفة السلسلة التالية من الأعداد، حدد المنوال.

أ ـ 50 و 45 و 55 و 55 و 45 و 50 و 55 و 45 و 55.

ب ـ 89 و 87 و 88 و 83 و 86 و 82 و 84.

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

٢٤ـ حدد خلية المنوال للبيانات فى كل مما يلى :

أ ـ المشكلة رقم ٤.

ب ـ المشكلة رقم ٥.

ج ـ المشكلة رقم ٦.

د ـ المشكلة رقم ٧.

هـ ـ المشكلة رقم ١٧.

و ـ المشكلة رقم ١٨.

٢٥ـ حدد المدى لكل فئة من فئات الأعداد التالية :

أ ـ 16 و 25 و 18 و 17 و 16 و 21و 14

ب ـ 45 و 39 و 42 و 42 و 43

ج ـ بيانات المشكلة رقم ٤

د ـ بيانات المشكلة رقم ٥

٢٦ـ 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؟

٢٧ـ كانت أربع قراءات لسمك 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؟.

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

نقطة مركز الخلية (%)	التكرار (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

٢٩ـ حدد الانحراف المعيارى للعينة لكل مما يلى :

أ ـ بيانات المشكلة رقم ٧.

ب ـ بيانات المشكلة رقم ١٧.

٣٠ـ 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

٣١ـ باستخدام بيانات المشكلة رقم ١٧ ، ارسم :

أ ـ مضلعات تكراريا.

ب ـ منحنى متجمعا صاعداً.

٣٢ـ باستخدام بيانات المشكلة رقم ١٨ ، ارسم :

أ ـ مضلعا تكراريا.

ب ـ منحنى متجمعا صاعدا.

٣٣ـ باستخدام بيانات المشكلة رقم ٢٨ ، ارسم :

أ ـ مضلعا تكراريا.

ب ـ منحنى متجمعا صاعدا.

٣٤ـ باستخدام بيانات المشكلة رقم ٣٠، ارسم :

أ ـ مضلعا تكراريا.

ب ـ منحنى متجمعا صاعدا.

٣٥ـ باستخدام بيانات المشكلة رقم ١٧ ، ارسم :

أ ـ مدرجا تكراريا.

ب ـ مدرجا تكراريا نسبيا.

ج ـ مدرجا تكراريامتجمعا صاعدا.

د ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

٣٦ـ باستخدام بيانات المشكلة رقم ١٨ ، ارسم :

أ ـ مدرجا تكراريا.

ب ـ مدرجا تكراريا نسبيا.

ج ـ مدرجا تكراريا متجمعا صاعدا.

د ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

٣٧ـ باستخدام بيانات المشكلة رقم ٢٨، ارسم :

أ ـ مدرجا تكراريا.

ب ـ مدرجا تكراريا نسبيا.

ج ـ مدرجا تكراريا متجمعا صاعدا.

د ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

٣٨ـ باستخدام بيانات المشكلة رقم ٣٠، ارسم :

أ ـ مدرجا تكراريا.

ب ـ مدرجا تكراريا نسبيا.

ج ـ مدرجا تكراريا متجمعا صاعدا.

د ـ مدرجا تكراريا متجمعا صاعدا نسبيا.

٣٩ـ حدد الالتواء والتفرطح لكل مما يلى :

أ ـ المشكلة رقم ٤.

ب ـ المشكلة رقم ٥.

ج ـ المشكلة رقم ٦.

د ـ المشكلة رقم ٧.

هـ ـ المشكلة رقم ١٨.

و ـ المشكلة رقم ٣٠.

٤٠ـ إذا كانت أقصى ضوضاء مسموح بها هى 134.5 ديسيبل، ما النسبة المئوية لبيانات المشكلة رقم ١٧ التى تقع أعلى من هذه القيمة؟

٤١ـ قوم المدرج التكرارى للمشكلة رقم ١٨، حيث تكون المواصفات هى: 0.60 + 4.25 كج.

٤٢ـ 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؟

٤٣ـ الوسط الحسابى للمجتمع لشركة دراجات سباق كان 9.07كج (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.

٤٤ـ 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.

٤٥ـ 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؟

٤٦ـ 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.

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

أ ـ الخاص بالمشكلة رقم ٤.

ب ـ الخاص بالمشكلة رقم ٥.

ج ـ الخاص بالمشكلة رقم ٦.

د ـ الخاص بالمشكلة رقم ٧.

هـ ـ الخاص بالمشكلة رقم ١٨.

و ـ الخاص بالمشكلة رقم ٣٠.

٤٨ـ باستخدام ورقة الاحتمالات الطبيعية، حدد (احكم على) الطبيعية لتوزيع كل مما يلى :

أ ـ العمود الثانى لجدول ٢ ـ ٤.

ب ـ أول ثلاثة أعمدة للمشكلة رقم ٥.

ج ـ العمود الثانى للمشكلة رقم ٦.

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

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

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