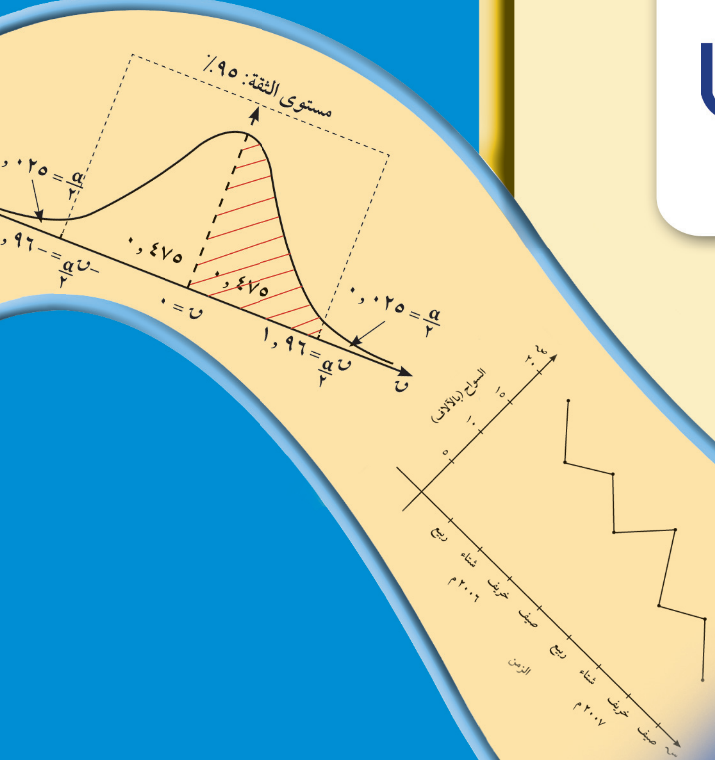
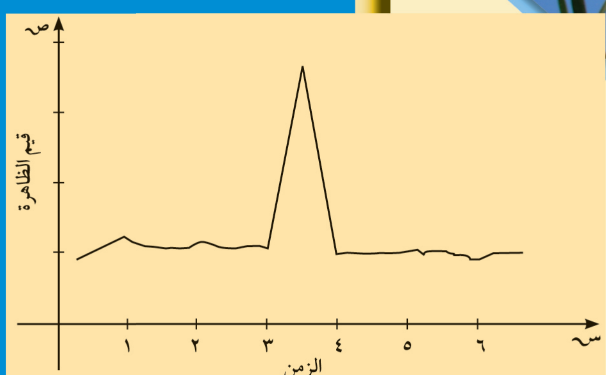
وزارة التربية

الرياضيات

كتاب الطالب





الصفّ الثاني عشر أدبي ١٢

الفصل الدراسي الأوّل

الطبعة الثانية

وزارة التربية

الرياضيات

الصفّ الثاني عشر أدبي
الفصل الدراسي الأول

كتاب الطالب

اللجنة الإشرافية لدراسة ومواءمة سلسلة كتب الرياضيات

أ. حسين علي عبدالله (رئيسًا)

أ. حصة يونس محمد علي

أ. فتحية محمود أبو زور

الطبعة الثانية

١٤٤١ - ١٤٤٢ هـ

٢٠٢٠ - ٢٠٢١ م

الطبعة الأولى ٢٠١٤م
الطبعة الثانية ٢٠١٦م
٢٠١٨م
٢٠٢٠م

فريق عمل دراسة ومواءمة كتب الرياضيات للصف الثاني عشر أدبي

أ. فتحي محمد عبد الفتاح (رئيسًا)

أ. محمود عبد الغني محمد أ. سعيد أحمد علي خلف
أ. يسرى شملان أحمد البحر أ. عيدة خلف عواد الشمري

أ. هنادي حباس غنيم المجول

دار التَّربَويّون House of Education ش.م.م. وبيرسون إديوكيشن ٢٠١٤م

الكتاب كاملًا

شاركنا بتقييم مناهجنا

ذات السلاسل - الكويت

أودع بمكتبة الوزارة تحت رقم (٤) بتاريخ ٣١/١/ ٢٠١٦م

حضرة صاحب السمو الشيخ نواف الأحمد الجابر الصباح
أمير دولة الكويت
H.H. Sheikh Nawaf AL-Ahmad Al-Jaber Al-Sabah
The Amir Of The State Of Kuwait

سمو الشيخ مشعل الأحمد الجابر الصباح
ولي عهد دولة الكويت
H.H. Sheikh Meshal AL-Ahmad Al-Jaber Al-Sabah
The Crown Prince Of The State Of Kuwait

مقدمـــــة

الحمدلله رب العالمين، والصلاة والسلام على سيد المرسلين، محمد بن عبدالله وصحبه أجمعين.

عندما شرعت وزارة التربية في عملية تطوير المناهج، استندت في ذلك إلى جملة من الأسس والمرتكزات العلمية والفنية والمهنية، حيث راعت متطلبات الدولة وارتباط ذلك بسوق العمل، وحاجات المتعلمين والتطور المعرفي والعلمي، بالإضافة إلى جملة من التحديات التي تمثلت بالتحدي القيمي والاجتماعي والاقتصادي والتكنولوجي وغيرها، وإن كنا ندرك أن هذه الجوانب لها صلة وثيقة بالنظام التعليمي بشكل عام وليس المناهج بشكل خاص.

ومما يجب التأكيد عليه، أن المنهج عبارة عن كم الخبرات التربوية والتعليمية التي تُقدم للمتعلم، وهذا يرتبط أيضًا بعمليات التخطيط والتنفيذ، والتي في محصلتها النهائية تأتي لتحقيق الأهداف التربوية، وعليه أصبحت عملية بناء المناهج الدراسية من أهم مكونات النظام التعليمي، لأنها تأتي في جانبين مهمين لقياس كفاءة النظام التعليمي، فهي من جهة تمثل أحد المدخلات الأساسية ومقياسًا أو معيارًا من معايير كفاءته من جهة أخرى، عدا أن المناهج تدخل في عملية إنماء شخصية المتعلم في جميع جوانبها الجسمية والعقلية والوجدانية والروحية والاجتماعية.

من جانب آخر، فنحن في قطاع البحوث التربوية والمناهج، عندما نبدأ في عملية تطوير المناهج الدراسية، ننطلق من كل الأسس والمرتكزات التي سبق ذكرها، بل إننا نراها محفزات واقعية تدفعنا لبذل قصارى جهدنا والمضي قدمًا في البحث في المستجدات التربوية سواء في شكل المناهج أم في مضامينها، وهذا ما قام به القطاع خلال السنوات الماضية، حيث البحث عن أفضل ما توصلت إليه عملية صناعة المناهج الدراسية، ومن ثم إعدادها وتأليفها وفق معايير عالمية استعدادًا لتطبيقها في البيئة التعليمية.

ولقد كانت مناهج العلوم والرياضيات من أول المناهج التي بدأنا بها عملية التطوير، إيمانًا بأهميتها وانطلاقًا من أنها ذات صفة عالمية، مع الأخذ بالحسبان خصوصية المجتمع الكويتي وبيئته المحلية، وعندما أدركنا أنها تتضمن جوانب عملية التعلم ونعني بذلك المعرفة والقيم والمهارات، قمنا بدراستها وجعلها تتوافق مع نظام التعليم في دولة الكويت، مركزين ليس فقط على الكتاب المقرر ولكن شمل ذلك طرائق وأساليب التدريس والبيئة التعليمية ودور المتعلم، مؤكدين على أهمية التكامل بين الجوانب العلمية والتطبيقية حتى تكون ذات طبيعة وظيفية مرتبطة بحياة المتعلم.

وفي ضوء ما سبق من معطيات وغيرها من الجوانب ذات الصفة التعليمية والتربوية تم اختيار سلسلة مناهج العلوم والرياضيات التي أكملناها بشكل ووقت مناسبين، ولنحقق نقلة نوعية في مناهج تلك المواد، وهذا كله تزامن مع عملية التقويم والقياس للأثر الذي تركته تلك المناهج، ومن ثم عمليات التعديل التي طرأت أثناء وبعد تنفيذها، مع التأكيد على الاستمرار في القياس المستمر والمتابعة الدائمة حتى تكون مناهجنا أكثر تفاعلية.

د. سعود هلال الحربي

الوكيل المساعد لقطاع البحوث التربوية والمناهج

المحتوَياتُ

الوحدة الأولى: التقدير واختبارات الفروض ١٠

(١-١) التقدير ١٢

(١-١-أ) التقدير بنقطة ١٣

(١-١-ب) التقدير بفترة الثقة ١٤

اولًا – إذا كان التباين للمجتمع σ^2 معلوم ١٨

ثانيًا – إذا كان التباين للمجتمع σ^2 غير معلوم، ن > ٣٠ .... ٢١

ثالثًا – إذا كان التباين للمجتمع σ^2 غير معلوم، ن ≤ ٣٠ .... ٢٢

(١-٢) اختبارات الفروض الإحصائية ٢٧

(١-٢-أ) إذا كان الانحراف المعياري للمجتمع σ معلوم ٢٩

(١-٢-ب) إذا كان الانحراف المعياري للمجتمع σ غير معلوم، ن > ٣٠ .... ٣١

(١-٢-جـ) إذا كان الانحراف المعياري للمجتمع σ غير معلوم، ن ≤ ٣٠ .... ٣٢

الوحدة الثانية: الارتباط والانحدار ٣٨

(٢-١) الارتباط ٤٠

(٢-١-أ) المخطط الانتشاري ٤١

(٢-١-ب) معامل الارتباط الخطي ٤٤

(٢-٢) الانحدار ٥٤

الوحدة الثالثة: السلاسل الزمنية ٦٤

(٣-١) السلسلة الزمنية ٦٦

(٣-٢) عناصر السلسلة الزمنية ٦٩

(٣-٣) تحليل السلسلة الزمنية ٧٧

– معادلة الاتجاه العام للسلسلة الزمنية ٧٧

الوحدة الأولى

التقدير واختبارات الفروض

Estimation and Hypotheses Testing

مشروع الوحدة: ما هي أفضل طريقة لإيجاد وظيفة؟

١ مقدمة المشروع: بعد التخرج يواجه الحاصلون على الإجازات والشهادات الجامعيّة تحدّ جديد هو الانخراط في سوق العمل.

٢ الهدف: هو البحث عن فرص عمل من خلال القيام بعدة خطوات ومحاولات متنوعة واستخدام العديد من الوسائل.

٣ اللوازم: حاسوب - شبكة الإنترنت.

٤ أسئلة حول التطبيق:

أ كيف ستختار عينة عشوائية من الموظفين للاستفسار عن الوسيلة التي استخدموها في إيجاد وظيفتهم؟

ب ما الخيارات التي اكتشفتها؟ نظّمها في استمارة.

(إرشاد):

- من خلال الأصدقاء والمعارف.
- من خلال الإعلانات في الصحف والمجلات.
- من خلال الوكالات المختصة في الربط بين سوق العمل وطالبي الوظائف.
- من خلال البحث عبر شبكة الإنترنت.
- من خلال التقدم مباشرة لطلب وظيفة من الشركة المختصة أو اعتماد وسيلة أخرى (اذكرها...).

جـ حدّد النسب المئوية لكل خيار ممّا سبق.

٥ التقرير: اكتب تقريرًا مفصلًا يحدّد النسب التي حصلت عليها من خلال العينة العشوائيّة التي اعتمدتها مكونًا جدولًا بالنسب المئوية عن كل وسيلة تمّ استخدامها لإيجاد وظيفة.

القرار: ضمّن تقريرك بعض الاقتراحات والنصائح والاستنتاجات التي نتجت عن تلك الدراسة.

دروس الوحدة

١-١ التقدير	١-٢ اختبارات الفروض الإحصائية
(١-١-أ) التقدير بنقطة	(١-٢-أ) σ معلومة
(١-١-ب) التقدير بفترة الثقة	(١-٢-ب) σ غير معلومة، ن > ٣٠
	(١-٢-جـ) σ غير معلومة، ن ≤ ٣٠

الوحدة الأولى

أضف إلى معلوماتك

في الوسائل الإعلامية المرئية والمسموعة والمكتوبة تطالعك نتائج إحصائية تتحدث عن توقعات أحداث معينة تتناول انتخابات نيابية أو رئاسية أو مبيعات أو مباريات... وأكثر ما يستوقفك هو نسبة مئوية معينة مع هامش خطأ محدد والسؤال المهم هو: كيف يتم التقدير وكيف يحتسب هامش الخطأ؟ توفر دروس هذه الوحدة فرصة أمام الطلاب للتعرّف على التقدير وهامش الخطأ والفروض الإحصائية وكيفية احتسابها.

أين أنت الآن (المعارف السابقة المكتسبة)

- تعلمت مقاييس النزعة المركزية: المتوسط الحسابي - الوسيط - المنوال.
- تعلمت المجتمع الإحصائي.
- تعلمت العينة واستخداماتها.

ماذا سوف تتعلم؟

- يُعرّف المعلمة والإحصاءة.
- إيجاد التقدير بنقطة.
- إيجاد التقدير بفترة ثقة.
- استكشاف الفروض الإحصائية.
- يُعرّف الاختبارات الإحصائية.
- اتخاذ القرار المناسب.

المصطلحات الأساسية

المعلمة - الإحصاءة - التقدير - التقدير بنقطة - فترة الثقة - الفروض الإحصائية - المقياس الإحصائي - فرض العدم - فرض البديل - القرار - مستوى المعنوية - درجات الحرية.

١ — ١

التقدير

Estimation

سوف تتعلم

- إيجاد التقدير بنقطة.
- إيجاد التقدير بفترة ثقة.

دعنا نفكر ونتناقش

متوسط درجات طلّاب الصف الثاني عشر في مادّة الرياضيات (حيث النهاية العظمى ١٠٠ درجة) في ٥ مدارس بالكويت $\overline{س}$ = ٨١

- هل يمكن استخدام هذه العيّنة لتقدير متوسّط الدرجات في كافة مدارس الكويت؟
- ما هي أفضل وسيلة للتقدير لنقترب من الحقيقة؟

سبق لنا في الصف الحادي عشر تعريف المجتمع الإحصائي والعينة العشوائية والأسباب التي تؤدي إلى أخذ العينات لدراسة المجتمع بدلًا من الحصر الشامل، وذلك لتقدير الوسط الحسابي للمجتمع μ أو الانحراف المعياري له σ.

ويعتبر الوسط الحسابي للمجتمع μ والانحراف المعياري للمجتمع σ من معالم المجتمع، وعادة ما تكون هذه المعالم مجهولة.

ولتقدير هذه المعالم نلجأ إلى سحب عينة عشوائية منه، ثم نحسب المتوسط الحسابي للعينة $\overline{س}$ أو الانحراف المعياري ع والذي يعتمد على قيم العينة ولا يعتمد على معالم المجتمع.

ملاحظة:
سنعتبر أن المجتمع الذي أخذت منه العينة يتبع التوزيع الطبيعي المعياري.

المعلمة (Parameter):
هي ثابت يصف المجتمع أو يصف توزيع المجتمع كالمتوسط الحسابي μ أو الانحراف المعياري σ.

الإحصاءة (Statistic Function):
هو اقتران تتعين قيمته من العينة كالمتوسط الحسابي $\overline{س}$ أو الانحراف المعياري ع.

تقدير المعلمة (Parameter Estimate):
هو إحصاءة تعتمد على قيم العينة وتعكس قيمة قريبة لمعلمة المجتمع ككلّ وتوزيعه.

في هذا الدرس سوف تتعرف طريقتين تساعدان على إيجاد قيم تقديرية لبعض معالم مجتمع معين:

- **طريقة أولى:** التقدير بنقطة.
- **طريقة ثانية:** التقدير بفترة الثقة.

(١–١–م) التقدير بنقطة **Point Estimate**

التقدير بنقطة هي قيمة وحيدة محسوبة من العينة تستخدم لتقدير معلمة مجهولة من معالم المجتمع.

فمثلًا المتوسط الحسابي للعينة العشوائية $\overline{س}$ يستخدم كتقدير بنقطة للمتوسط الحسابي للمجتمع μ، وكذلك الانحراف المعياري للعينة ع يستخدم كتقدير بنقطة للانحراف المعياري للمجتمع σ.

مثال (١)

تبيّن البيانات التالية معدّل درجة الحرارة عند ٤٠ شخصًا بحالة صحية جيدة:

٣٧	٣٧	٣٦,٧	٣٦,٧	٣٧,٢	٣٦,٩	٣٦,٩	٣٦,٩	٣٦,٩	٣٧,٤
٣٧,١	٣٧	٣٦,١	٣٦,١	٣٧,١	٣٦,٤	٣٦,٥	٣٧,١	٣٦,٦	٣٦,٦
٣٦,٨	٣٦,٩	٣٦,٤	٣٧	٣٦,٣	٣٧,٢	٣٧	٣٧,٥	٣٦,٤	٣٦,٣
٣٦,٤	٣٦,٨	٣٧,٥	٣٧,١	٣٧,٤	٣٦,٨	٣٦,٧	٣٧	٣٧	٣٦,٢

استخدم هذه العينة لقيم معدل درجة الحرارة لتوجد أفضل تقدير بنقطة للمتوسط الحسابي μ لمعدل درجة حرارة مجتمع أخذت منه هذه العينة.

الحل:

نوجد المتوسط الحسابي $\overline{س}$ لقيم البيانات في العينة التي تمثل معدل درجة الحرارة عند ٤٠ شخصًا بحالة صحية جيدة.

نوجد المتوسط الحسابي $\overline{س}$ لقيم البيانات في العينة التي تمثل معدل درجة الحرارة عند ٤٠ شخصًا بحالة صحية جيدة.

$$\overline{س} = \frac{\sum س_ر}{ن}$$

$$= \frac{١٤٧٢,٨}{٤٠} = ٣٦,٨٢°$$

$\therefore$ القيمة التقديرية للمتوسط الحسابي μ لمعدل درجة حرارة المجتمع الذي أخذت منه هذه البيانات هي μ = ٣٦,٨٢°

حاول أن تحل

١ تبيّن البيانات التالية درجات ٤٠ طالبًا في مادة الرياضيات حيث النهاية العظمى ٢٠ درجة.

١٦، ١٥، ١٤، ١٧، ١٨، ١٩، ١٧، ١٥، ١٤، ١٣، ١٢، ١٣، ٩، ١٠، ١٢، ١٤، ٨، ١٦، ١٩، ٧،
١٤، ١٥، ١٦، ١٨، ١٠، ١١، ٩، ٨، ١٥، ١٢، ١٩، ١٤، ١٠، ١١، ١٥، ١٦، ١٤، ١٧، ١٨، ١٦.

استخدم هذه العينة لقيم الدرجات لتوجد التقدير بنقطة للمتوسط الحسابي للمجتمع μ الذي أخذت منه هذه العينة.

(١–١–ب) التقدير بفترة الثقة Confidence Interval Estimation

علمنا مما سبق أن لكل مجتمع معالم منها المتوسط الحسابي μ والانحراف المعياري σ، ودرسنا كيفية إيجاد التقدير بنقطة لتلك المعالم. وعلمنا أن التقدير بنقطة لإحدى معالم المجتمع هو قيمة وحيدة محسوبة من العينة وبالتالي فإن احتمال الخطأ في التقدير بنقطة يكون كبيرًا. ولذلك فإنه من الأفضل إيجاد فترة معينة يتوقع أن تقع معلمة المجتمع داخلها بنسبة معينة أو باحتمال معين. إن مثل هذه الفترة تسمى **فترة الثقة**.

فترة الثقة Confidence Interval

تعريف: فترة الثقة

هي فترة طرفاها متغيران عشوائيان (أي أنها فترة عشوائية) تستخدم لتقدير إحدى معالم المجتمع.

وهذه الفترة تحوي إحدى معالم المجتمع بنسبة معينة تسمى مستوى الثقة، فمثلًا إذا كان مستوى الثقة ٩٥٪ فإن نسبة الخطأ في التقدير تكون ٥٪

يرمز لمستوى الثقة بالرمز ١٠٠٪ (١ – α) حيث (١ – α) هو معامل مستوى الثقة وα هي نسبة الخطأ في التقدير.

وعلى سبيل المثال:

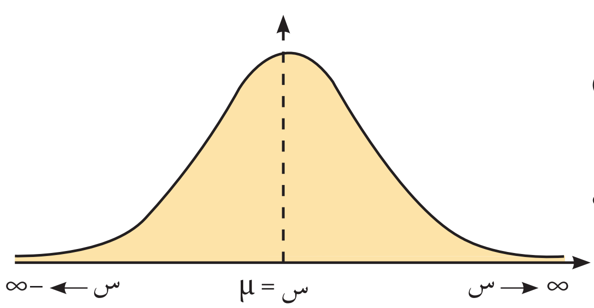
- إذا كان مستوى الثقة ٩٠٪ فإن مستوى المعنوية α = ٠,١٠
- وإذا كان مستوى الثقة ٩٥٪ فإن مستوى المعنوية α = ٠,٠٥
- أيضًا إذا كان مستوى الثقة ٩٩٪ فإن مستوى المعنوية α = ٠,٠١

ومن هذه الخيارات الثلاثة، يعتبر مستوى الثقة ٩٥٪ هو الأكثر انتشارًا لأنه يؤمن التوازن الأنسب بين الدقة الموضحة من خلال طول فترة الثقة والدقة الموضحة من خلال مستوى الثقة.

منحنى التوزيع الطبيعي Curve of Normal Distribution

تعرفنا فيما سبق على بيان منحنى التوزيع الطبيعي، وعلمنا من خواص التوزيع الطبيعي ما يلي:

- المتوسط الحسابي = الوسيط = المنوال.
- يكون بيان المنحنى على شكل ناقوس (جرس) متماثل حول محوره (س = μ).
- يمتد المنحنى من طرفيه إلى +∞ وإلى –∞ (لا يقطع المحور الأفقي).
- المساحة تحت المنحنى تساوي الواحد الصحيح (وحدة مساحة).



المستقيم الرأسي س = μ يقسم المساحة تحت المنحنى إلى منطقتين متماثلتين مساحة كل منهما تساوي نصف وحدة مساحة كما في الشكل.

منحنى التوزيع الطبيعي المعياري
Curve of Standard Normal Distribution

إذا كان المتوسط الحسابي للتوزيع الطبيعي μ = صفر والانحراف المعياري σ = ١

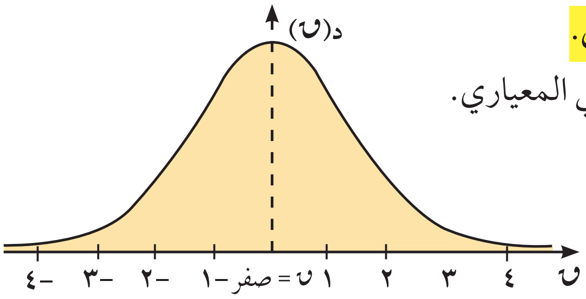
يسمى التوزيع الطبيعي **بالتوزيع الطبيعي المعياري**.

الشكل المرسوم يمثّل بيان منحنى التوزيع الطبيعي المعياري.

المستقيم ں = صفر هو محور التماثل للمنحنى.

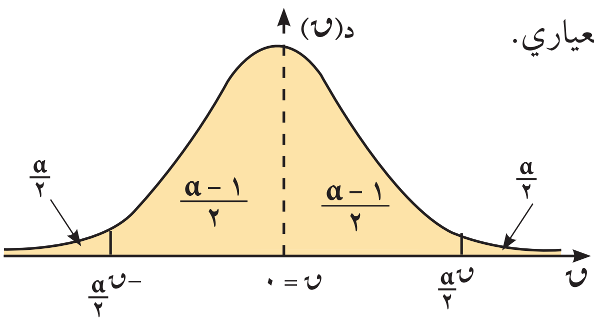
تأخذ ں قيم موجبة وتزداد جهة اليمين بينما

تأخذ ں قيمًا سالبة وتنقص جهة اليسار.



القيمة الحرجة
Critical Value

الشكل المرسوم يبيّن منحنى التوزيع الطبيعي المعياري.



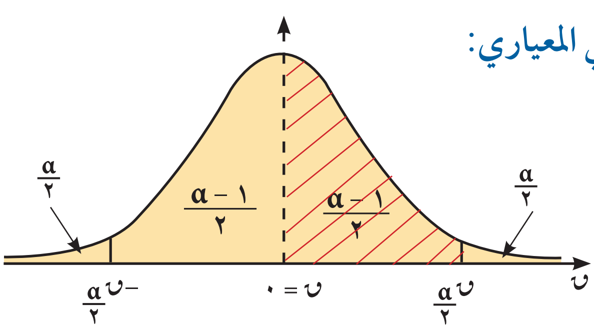
- نعلم أن المساحة تحت المنحنى الطبيعي تساوي الواحد (وحدة المساحة) ولتمثيل (١ − α) من المساحة الكلية تحت منحنى التوزيع الطبيعي المعياري نحصر هذه المساحة بين حدين رأسيين متساويي البعد عن المحور الرأسي كما هو موضّح في الشكل.

نلاحظ أن المحور الرأسي يقسم المساحة (١ − α) إلى نصفين كل منهما يساوي $\frac{١ - \alpha}{٢}$.

تكون المساحة المتبقية من المساحة الكلية هي α موزعة على طرفي المنحنى بالتساوي كل منها يساوي $\frac{\alpha}{٢}$.

- نعبّر عن الحدين الرأسيين بالرمز $\text{ں}_{\frac{\alpha}{2}} = \text{ں}_{\frac{١ - \alpha}{٢}}$ وبالرمز $-\text{ں}_{\frac{\alpha}{2}} = -\text{ں}_{\frac{١ - \alpha}{٢}}$، حيث $\text{ں}_{\frac{\alpha}{2}}$ يفصل مساحة $\frac{\alpha}{٢}$ من ذيل الطرف الأيمن ومساحة $\frac{١ - \alpha}{٢}$ من المستقيم ں = صفر، بينما $-\text{ں}_{\frac{\alpha}{2}}$ يفصل مساحة $\frac{\alpha}{٢}$ من ذيل الطرف الأيسر ومساحة $\frac{١ - \alpha}{٢}$ من المستقيم ں = صفر.
- تسمى القيمة الموجبة $\text{ں}_{\frac{\alpha}{2}}$ **بالقيمة الحرجة (Critical Value)**.

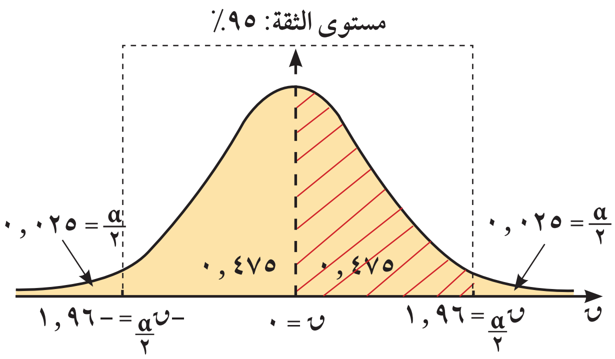
إيجاد القيمة الحرجة من جدول التوزيع الطبيعي المعياري:



لإيجاد قيمة $\text{ں}_{\frac{\alpha}{2}}$ المناظرة للمساحة تحت المنحنى نحسب المساحة $\frac{١ - \alpha}{٢}$ التي تقع على يسار $\text{ں}_{\frac{\alpha}{2}}$ ويمين الصفر أي في الفترة $[٠ ، \text{ں}_{\frac{\alpha}{2}}]$ ثم نكشف عنها في الجدول المرفق في نهاية الوحدة حيث العمود الأول قيم ں ابتداءً من ٠,٠ وحتى ٣,١٠ وأكثر. والصف الأول يمثل الأجزاء من المئة لقيم ں، ومنه يمكن تحديد قيمة $\text{ں}_{\frac{\alpha}{2}}$.

مثال (٢)

أوجد القيمة الحرجة $ن_{\frac{\alpha}{٢}}$ المناظرة لمستوى ثقة ٩٥٪ باستخدام جدول التوزيع الطبيعي المعياري.



الحل:

$\because$ مستوى الثقة هو ٩٥٪

$\therefore$ $١ - \alpha = ٠,٩٥$

$\therefore$ $\frac{١ - \alpha}{٢} = \frac{٠,٩٥}{٢} = ٠,٤٧٥$

نبحث في جدول التوزيع الطبيعي المعياري عن قيمة ن المناظرة للعدد ٠,٤٧٥٠

فنجد $ن_{\frac{\alpha}{٢}} = ن_{٠,٤٧٥٠} = ١,٩٦$

حاول أن تحل

٢ أوجد القيمة الحرجة $ن_{\frac{\alpha}{٢}}$ المناظرة لمستوى ثقة ٩٧٪ باستخدام جدول التوزيع الطبيعي المعياري.

مثال (٣)

أوجد القيمة الحرجة $ن_{\frac{\alpha}{٢}}$ المناظرة لمستوى ثقة ٩٠٪ باستخدام جدول التوزيع الطبيعي المعياري.

الحل:

$\because$ مستوى الثقة هو ٩٠٪

$\therefore$ $١ - \alpha = ٠,٩٠$

$\therefore$ $\frac{١ - \alpha}{٢} = \frac{٠,٩٠}{٢}$

$= ٠,٤٥$

نبحث في الجدول عن القيمة ٠,٤٥٠٠ فنجدها تقع بين القيمتين ٠,٤٤٩٥ ، ٠,٤٥٠٥

أي أن $ن_{\frac{\alpha}{٢}}$ تقع بين ١,٦٤ ، ١,٦٥

لذا نأخذ المتوسط الحسابي للقيمتين ١,٦٤ ، ١,٦٥ كتقدير لقيمة $ن_{\frac{\alpha}{٢}}$

$\therefore$ $ن_{\frac{\alpha}{٢}} = \frac{١,٦٤ + ١,٦٥}{٢} = \frac{٣,٢٩}{٢}$

$\therefore$ $ن_{\frac{\alpha}{٢}} = ١,٦٤٥$

حاول أن تحل

٣ أوجد القيمة الحرجة $ن_{\frac{\alpha}{٢}}$ المناظرة لمستوى ثقة ٩٩٪، باستخدام جدول التوزيع الطبيعي المعياري.

هامش الخطأ Margin of Error

أوّلًا: الخطأ بالتقدير بنقطة Point Estimation Error

علمنا فيما سبق أنه يمكن استخدام المتوسط الحسابي للعينة $\bar{\text{س}}$ كتقدير بنقطة للمتوسط الحسابي للمجتمع μ.

ومن المتوقع أن تكون قيمة المتوسط الحسابي للعينة $\bar{\text{س}}$ غير مساوية لقيمة المتوسط الحسابي للمجتمع μ.

تسمى القيمة المطلقة للفرق بين القيمتين السابقتين بالخطأ المعياري

وتساوي $\frac{\sigma}{\sqrt{\text{ن}}}$ حيث σ الانحراف المعياري للمجتمع، ن عدد قيم العينة (أو حجم العينة).

ثانيًا: الخطأ بالتقدير بفترة Interval Estimation Error

والآن نتعرض للخطأ بالتقدير بفترة فعندما نستخدم عينة لتقدير المتوسط الحسابي لمجتمع μ، يكون الخطأ في التقدير هو القيمة المطلقة للفرق بين المتوسط الحسابي للعينة $\bar{\text{س}}$، والمتوسط الحسابي للمجتمع μ ويعرف هامش الخطأ هـ :

$\text{هـ} = \text{ى}_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{\text{ن}}}$ باحتمال $(1 - \alpha)$، حيث α تعبر عن نسبة الخطأ في التقدير.

وحتى يكون هامش الخطأ أقل ما يمكن يجب أن تتحقق المتباينة:

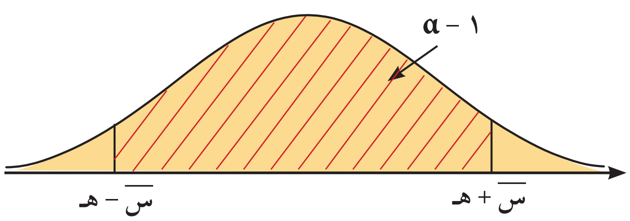
$$|\bar{\text{س}} - \mu| < \text{هـ}$$

أي أن: $$|\mu - \bar{\text{س}}| < \text{هـ}$$

$$-\text{هـ} < \mu - \bar{\text{س}} < \text{هـ}$$

$$\bar{\text{س}} - \text{هـ} < \mu < \bar{\text{س}} + \text{هـ}$$

وعليه تكون فترة الثقة هي $(\bar{\text{س}} - \text{هـ} \ ، \ \bar{\text{س}} + \text{هـ})$



التقدير بفترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ Confidence Interval Estimation for the Mean Value μ of Statistical Population

أوّلًا: إذا كان التباين للمجتمع σ^2 معلوم

إذا أخذت عينة عشوائية حجمها ن من مجتمع طبيعي ط$(\mu، \sigma^2)$ وتباينه σ^2 معلوم فإن تقدير فترة الثقة ١٠٠٪ $(١ - \alpha)$ للمتوسط الحسابي μ هي: $(\overline{س} - هـ ، \overline{س} + هـ)$

حيث $\overline{س}$ المتوسط الحسابي للعينة، هـ هامش الخطأ.

وتسمى القيمتان $\overline{س}$ − هـ ، $\overline{س}$ + هـ **طرفي فترة الثقة**.

ملاحظة: عند إيجاد فترة الثقة ١٠٠٪ $(١ - \alpha)$ سنكتفي بمستوى الثقة ٩٥٪ والتي تناظرها القيمة الحرجة $ى_{\frac{\alpha}{٢}} = ١,٩٦$ (من جدول التوزيع الطبيعي المعياري).

تفسير فترة الثقة

عند اختيار عينات عشوائية مختلفة متساوية في الحجم (ن) وحساب حدود فترة الثقة لكل عينة فإننا نتوقع أن ٩٥٪ من فترات الثقة هذه تحوي القيمة الحقيقية للمتوسط الحسابي للمجتمع (μ).

فمثلًا عند اختيار ١٠٠ عينة عشوائية ذات الحجم نفسه (ن) وفي كل مرة نحسب $\overline{س}$ وفترة الثقة فإننا نتوقع أن ٩٥ فترة تحوي μ الحقيقية و٥ فترات لا تحويها.

الخطوات المتبعة لإيجاد فترة الثقة للمتوسط الحسابي μ

إذا كانت σ^2 معلومة حيث ن > ٣٠ أو ن ≤ ٣٠

1. نوجد القيمة الحرجة $ى_{\frac{\alpha}{٢}}$ المناظرة لمستوى ثقة ٩٥٪ وهي ١,٩٦ .
2. نوجد هامش الخطأ هـ = $ى_{\frac{\alpha}{٢}} \times \frac{\sigma}{\sqrt{ن}}$
3. نوجد فترة الثقة $(\overline{س} - هـ ، \overline{س} + هـ)$.

مثال (٤)

أجريت دراسة لعينة من الإناث حول معدل النبض لديهن فإذا كان حجم عينة الإناث ن = ٤٠ والانحراف المعياري لمجتمع الإناث σ = ١٢,٥ والمتوسط الحسابي للعينة $\overline{\text{س}}$ = ٧٦,٣. باستخدام مستوى ثقة ٩٥٪.

١ أوجد هامش الخطأ.

٢ أوجد فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

٣ فسّر فترة الثقة.

الحل:

١ ∵ مستوى الثقة ٩٥٪ ∴ القيمة الحرجة $\text{ى}_{\frac{\alpha}{2}}$ = ١,٩٦

بما أن σ معلومة ∴ هامش الخطأ هـ = $\text{ى}_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{\text{ن}}}$

∵ ن = ٤٠ ، σ = ١٢,٥ ، $\overline{\text{س}}$ = ٧٦,٣

∴ هـ = ١,٩٦ × $\frac{١٢,٥}{\sqrt{٤٠}}$

هـ ≈ ٣,٨٧٣٨

٢ فترة الثقة هي ($\overline{\text{س}}$ − هـ ، $\overline{\text{س}}$ + هـ)

= (٧٦,٣ − ٣,٨٧٣٨ ، ٧٦,٣ + ٣,٨٧٣٨)

= (٧٢,٤٢٦٢ ، ٨٠,١٧٣٨)

٣ عند اختيار ١٠٠ عينة عشوائية ذات الحجم نفسه (ن = ٤٠) وحساب حدود فترة الثقة لكل عينة فإننا نتوقّع أن ٩٥ فترة تحوي القيمة الحقيقية للمتوسط الحسابي للمجتمع μ.

حاول أن تحل

٤ من المثال (٤)، إذا أجريت الدراسة على عينة أخرى من الإناث حجمها ١٠٠ والانحراف المعياري لمجتمع الإناث σ = ٣,٦ والمتوسط الحسابي للعينة $\overline{\text{س}}$ = ١٨,٤ باستخدام مستوى ثقة ٩٥٪

١ أوجد هامش الخطأ.

٢ أوجد فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

٣ فسّر فترة الثقة.

مثال (٥)

أجريت دراسة لعينة من ١٨ طالبًا حول متوسط عدد ساعات استخدام الألواح الذكية (TABLETS) أسبوعيًّا. فإذا كان الانحراف المعياري σ = ١,٨ والمتوسط الحسابي للعينة $\overline{س}$ = ١٥، باستخدام مستوى ثقة ٩٥٪

١ أوجد هامش الخطأ.

٢ أوجد فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

٣ فسّر فترة الثقة.

الحل:

١ $\because$ مستوى الثقة ٩٥٪ $\therefore$ القيمة الحرجة $ى_{\frac{\alpha}{2}}$ = ١,٩٦

$\because$ σ معلومة $\therefore$ هامش الخطأ هـ = $ى_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{ن}}$

$\because$ ن = ١٨ ، σ = ١,٨ ، $\overline{س}$ = ١٥

$\therefore$ هـ = ١,٩٦ $\times \frac{١,٨}{\sqrt{١٨}}$

هـ $\approx$ ٠,٨٣١٦

٢ فترة الثقة هي ($\overline{س}$ − هـ ، $\overline{س}$ + هـ)

= (١٥ − ٠,٨٣١٦ ، ١٥ + ٠,٨٣١٦)

= (١٤,١٦٨٤ ، ١٥,٨٣١٦)

٣ عند اختيار ١٠٠ عينة عشوائية ذات الحجم نفسه (ن = ١٨) وحساب حدود فترة الثقة لكل عينة فإننا نتوقّع أن ٩٥ فترة تحوي القيمة الحقيقية للمتوسط الحسابي للمجتمع μ.

حاول أن تحل

٥ أجريت دراسة لعينة من ٢٤ طالبًا حول متوسط عدد ساعات مشاهدة التلفزيون أسبوعيًّا. فإذا كان الانحراف المعياري σ = ٢,٥ والمتوسط الحسابي للعينة $\overline{س}$ = ٢١، باستخدام مستوى ثقة ٩٥٪

١ أوجد هامش الخطأ.

٢ أوجد فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

٣ فسّر فترة الثقة.

ثانيًا: إذا كان التباين للمجتمع σ^2 غير معلوم وحجم العينة ن > ٣٠

الخطوات المتبعة لإيجاد فترة الثقة للمتوسط الحسابي μ

إذا كانت σ^2 غير معلومة حيث ن > ٣٠

١ نوجد القيمة الحرجة $ى_{\frac{\alpha}{2}}$ المناظرة لمستوى ثقة ٩٥٪ وهي ١,٩٦

٢ نوجد هامش الخطأ هـ = $ى_{\frac{\alpha}{2}} \times \frac{ع_ر}{\sqrt{ن}}$

٣ نوجد فترة الثقة ($\overline{س}$ − هـ ، $\overline{س}$ + هـ).

مثال (٦)

عينة عشوائية حجمها ٣٦ ، فإذا كان المتوسط الحسابي للعينة ٦٠ وتباينها ١٦ ، باستخدام مستوى ثقة ٩٥٪

١ أوجد هامش الخطأ.

٢ أوجد فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

٣ فسّر فترة الثقة.

الحل:

١ ∵ مستوى الثقة ٩٥٪ ∴ القيمة الحرجة $ى_{\frac{\alpha}{2}}$ = ١,٩٦

∵ σ غير معلوم ، ن > ٣٠ ∴ هامش الخطأ هـ = $ى_{\frac{\alpha}{2}} \times \frac{ع_ر}{\sqrt{ن}}$

∵ التباين $ع_ر^2$ = ١٦

∴ الانحراف المعياري $ع_ر$ = ٤

$\overline{س}$ = ٦٠ ، ن = ٣٦

هـ = ١,٩٦ × $\frac{٤}{\sqrt{٣٦}}$

≈ ١,٣٠٦٦

٢ فترة الثقة هي ($\overline{س}$ − هـ ، $\overline{س}$ + هـ)

= (٦٠ − ١,٣٠٦٦ ، ٦٠ + ١,٣٠٦٦)

= (٥٨,٦٩٣٤ ، ٦١,٣٠٦٦)

٣ عند اختيار ١٠٠ عينة عشوائية ذات الحجم نفسه (ن = ٣٦) وحساب حدود فترة الثقة لكل عينة فإننا نتوقّع أن ٩٥ فترة تحوي القيمة الحقيقية للمتوسط الحسابي للمجتمع μ.

حاول أن تحل

٦ أخذت عينة عشوائية من مجتمع طبيعي حجمها ن = ٨١ ومتوسطها الحسابي $\overline{س}$ = ٥٠، وانحرافها المعياري ع = ٩، باستخدام مستوى ثقة ٩٥٪

١ أوجد هامش الخطأ.

٢ أوجد فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

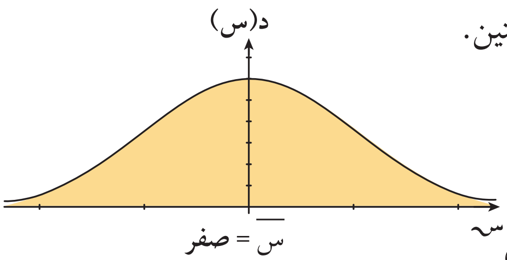
٣ فسّر فترة الثقة.

ثالثًا: إذا كان التباين للمجتمع σ^2 غير معلوم وحجم العينة ن ≤ ٣٠.

إذا أخذت عينة عشوائية حجمها ن من مجتمع طبيعي تباينه σ^2 غير معلوم وحجم العينة ن ≤ ٣٠ فإن توزيع العينة لا يؤول إلى التوزيع الطبيعي وفي هذه الحالة يلزمنا استخدام توزيع آخر هو توزيع ت للعينات الصغيرة التي حجمها ن ≤ ٣٠ ويكون تقدير فترة الثقة ١٠٠٪ (١ − α) للمتوسط الحسابي μ هي ($\overline{س}$ − هـ ، $\overline{س}$ + هـ) حيث $\overline{س}$ المتوسط الحسابي للعينة، هـ هامش الخطأ.

خواص التوزيع ت — Properties of t Distribution

١ توزيع متماثل حول متوسطه الحسابي والذي يساوي صفرًا، ويمتد إلى ∞ من جهة اليمين وإلى −∞ من جهة اليسار ويزداد قربًا من الصفر في الجهتين.



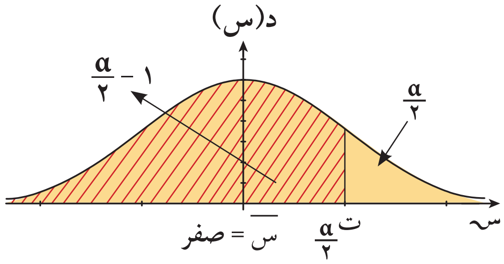
٢ انحرافه المعياري أكبر من الواحد.

٣ يعتمد هذا التوزيع على درجات الحرية والتي تساوي (حجم العينة − ١) أي (ن − ١).

٤ التوزيع ت يشبه التوزيع الطبيعي إلّا أن قمته أكثر انخفاضًا من التوزيع الطبيعي.

٥ كلما زادت درجات الحرية اقترب هذا التوزيع من التوزيع الطبيعي ويقترب انحرافه المعياري إلى الواحد الصحيح.

إيجاد القيمة الحرجة من جدول توزيع ت.



- لإيجاد القيمة الحرجة من جدول توزيع ت حيث يبيّن العمود الأول قيم درجات الحرية (ن – ١) وتبدأ من ١ إلى ٣٠ وأكثر والصف الأول يمثّل قيم $\frac{\alpha}{٢}$ ومنه يمكن تحديد $ت_{\frac{\alpha}{٢}} = ت_{١ - \frac{\alpha}{٢}}$.

مثال (٧)

أخذت عينة عشوائية بسيطة حجمها ن = ٢٣ من مجتمع طبيعي.

أوجد القيمة الحرجة $ت_{\frac{\alpha}{٢}}$ المناظرة لمستوى الثقة ٩٥٪ باستخدام جدول التوزيع ت.

الحل:

∵ ن = ٢٣

∴ درجات الحرية (ن – ١) = ٢٣ – ١

= ٢٢

∵ مستوى الثقة هو ٩٥٪

∴ ١ – α = ٠,٩٥

α = ٠,٠٥

$\frac{\alpha}{٢}$ = ٠,٠٢٥

ومن جدول التوزيع ت

تكون قيمة $ت_{\frac{\alpha}{٢}} = ت_{٠,٠٢٥}$ = ٢,٠٧٤

حاول أن تحل

٧ أخذت عينة عشوائية بسيطة حجمها ن = ٢٠ من مجتمع طبيعي.

أوجد القيمة الحرجة $ت_{\frac{\alpha}{٢}}$ المناظرة لمستوى الثقة ٩٥٪ باستخدام جدول التوزيع ت.

والآن، بعد أن علمنا كيف نوجد القيم الحرجة $ت_{\frac{\alpha}{٢}}$، يمكننا أن نوجد هامش الخطأ هـ وفترة الثقة.

هامش الخطأ للمتوسط الحسابي للمجتمع الإحصائي μ (في حالة $\sigma^{٢}$ غير معلوم، ن ≥ ٣٠)

Margin of Error for Mean Value of Statistical Population Where σ^2 is not known and $n \geq 30$

$هـ = ت_{\frac{\alpha}{٢}} \times \frac{ع}{\sqrt{ن}}$ حيث ع الانحراف المعياري للعينة

فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ (في حالة $\sigma^{٢}$ غير معلوم، ن ≥ ٣٠)

Confidence Interval for Mean Value of Statistical Population where σ^2 is not known and $n \leq 30$

$(\overline{س} - هـ ، \overline{س} + هـ)$

الخطوات المتبعة لإيجاد فترة الثقة للمتوسط الحسابي μ

إذا كانت $\sigma^{٢}$ غير معلومة، ن ≥ ٣٠

١ نوجد درجات الحرية (ن − ١).

٢ نوجد القيمة الحرجة $ت_{\frac{\alpha}{٢}}$ المناظرة لدرجة ثقة ٩٥٪ من جدول توزيع ت.

٣ نوجد هامش الخطأ $هـ = ت_{\frac{\alpha}{٢}} \times \frac{ع}{\sqrt{ن}}$

٤ نوجد فترة الثقة $(\overline{س} - هـ ، \overline{س} + هـ)$.

مثال (٨)

أخذت عينة عشوائية من مجتمع طبيعي حجمها ن = ٢٥، فإذا كان الانحراف المعياري للعينة ($ع_ر$) يساوي ١٠ ومتوسطها الحسابي ($\overline{س}$) يساوي ١٥، استخدم مستوى ثقة ٩٥٪ لإيجاد:

١ هامش الخطأ.

٢ فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

الحل:

١ ∵ $\sigma^٢$ غير معلوم ، ن ≤ ٣٠

∴ نستخدم توزيع ت.

∵ ن = ٢٥

∴ درجات الحرية (ن − ١) = ٢٥ − ١ = ٢٤

∵ مستوى الثقة ١ − α = ٩٥٪

∴ ١ − α = ٠,٩٥ ⟸ α = ٠,٠٥٠

∴ $\frac{\alpha}{٢}$ = ٠,٠٢٥

من جدول توزيع ت تكون قيمة $ت_{\frac{\alpha}{٢}}$ = $ت_{٠,٠٢٥}$ = ٢,٠٦٤

هامش الخطأ هـ = $ت_{\frac{\alpha}{٢}} \times \frac{ع_ر}{\sqrt{ن}}$

هـ = ٢,٠٦٤ × $\frac{١٠}{\sqrt{٢٥}}$

= ٤,١٢٨

٢ فترة الثقة = ($\overline{س}$ − هـ ، $\overline{س}$ + هـ)

= (١٥ − ٤,١٢٨ ، ١٥ + ٤,١٢٨)

= (١٠,٨٧٢ ، ١٩,١٢٨)

حاول أن تحل

٨ أوجد فترة ثقة ٩٥٪ للمتوسط الحسابي للمجتمع الإحصائي μ علمًا أن العينة أخذت من مجتمع طبيعي.

إذا كان لدينا $\overline{س}$ = ٨,٤ ، $ع_ر$ = ٢,٣ ، ن = ١٣

ويمكن تلخيص الحالات الثلاث السابقة كما في الجدول التالي:

الانحراف المعياري (σ)	حجم العينة (ن)	هامش الخطأ (هـ)	فترة الثقة ($\overline{\text{س}}$ − هـ ، $\overline{\text{س}}$ + هـ)
معلوم	ن > ٣٠ أو ن ≤ ٣٠	$\text{هـ} = \text{ى}_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{\text{ن}}}$	$(\overline{\text{س}} - \text{ى}_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{\text{ن}}} ، \overline{\text{س}} + \text{ى}_{\frac{\alpha}{2}} \times \frac{\sigma}{\sqrt{\text{ن}}})$
غير معلوم (نستبدل σ بـ ع)	ن > ٣٠	$\text{هـ} = \text{ى}_{\frac{\alpha}{2}} \times \frac{\text{ع}}{\sqrt{\text{ن}}}$	$(\overline{\text{س}} - \text{ى}_{\frac{\alpha}{2}} \times \frac{\text{ع}}{\sqrt{\text{ن}}} ، \overline{\text{س}} + \text{ى}_{\frac{\alpha}{2}} \times \frac{\text{ع}}{\sqrt{\text{ن}}})$
	ن ≤ ٣٠	$\text{هـ} = \text{ت}_{\frac{\alpha}{2}} \times \frac{\text{ع}}{\sqrt{\text{ن}}}$	$(\overline{\text{س}} - \text{ت}_{\frac{\alpha}{2}} \times \frac{\text{ع}}{\sqrt{\text{ن}}} ، \overline{\text{س}} + \text{ت}_{\frac{\alpha}{2}} \times \frac{\text{ع}}{\sqrt{\text{ن}}})$

مثال (٩)

أخذت عينة عشوائية حجمها ن = ٦٠، فإذا كان الانحراف المعياري للعينة (ع) يساوي ١٨ ومتوسطها الحسابي ($\overline{\text{س}}$) يساوي ٣٦ ، استخدم مستوى ثقة ٩٥٪ لإيجاد:

١ هامش الخطأ.

٢ فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

الحل:

١ ∵ σ^2 غير معلوم، ن ≥ ٣٠

∴ القيمة الحرجة $\text{ى}_{\frac{\alpha}{2}}$ المناظرة لمستوى ثقة ٩٥٪ = ١,٩٦ .

∵ ع = ١٨، ن = ٦٠

∴ $\text{هـ} = \text{ى}_{\frac{\alpha}{2}} \times \frac{\text{ع}}{\sqrt{\text{ن}}}$

$\text{هـ} = ١,٩٦ \times \frac{١٨}{\sqrt{٦٠}}$

$\approx ٤,٥٥٤٦$

٢ فترة الثقة = ($\overline{\text{س}}$ − هـ ، $\overline{\text{س}}$ + هـ)

= (٣٦ − ٤,٥٥٤٦ ، ٣٦ + ٤,٥٥٤٦)

= (٣١,٤٤٥٤ ، ٤٠,٥٥٤٦)

حاول أن تحل

٩ أخذت عينة عشوائية من ٢٠ نبتة لدراسة نموها. فإذا كان متوسط النمو = ٣٦ سم خلال عام والانحراف المعياري للعينة ٤,٦ سم، استخدم مستوى ثقة ٩٥٪ لإيجاد:

١ هامش الخطأ.

٢ فترة الثقة للمتوسط الحسابي للمجتمع الإحصائي μ.

١ — ٢

سوف تتعلم

- القيمة الحرجة.
- مستوى المعنوية.
- درجات الحرية.
- الفروض.
- اختبار الفروض.
- فرض العدم.
- الفرض البديل.

اختبارات الفروض الإحصائية

Satistical Hypotheses Testing

دعنا نفكر ونتناقش

يتتج مصنع نوعًا معينًا من المعلبات مسجّل على العلبة أن الوزن الصافي ٢٠٠ جرام.

فإذا تمّ أخذ عينة حجمها ١٠٠ علبة وتمّ حساب المتوسط الحسابي لأوزان هذه العينة فوجد أنه ١٩٧٫٣ جرامًا،

فهل يمكن الحكم على هذا المصنع بأنه يقوم بغش تجاري؟

ما هي حيثيات هذا الحكم؟

نحن نعلم أنه في كثير من الأحيان وفي مواقف معينة نحتاج إلى اتخاذ قرار بناء على معلومات محددة وحيثيات معقولة لها مبررها، لذلك دعت الضرورة إلى دراسة ما يسمى بالفرض الإحصائي واختبارات الفروض الإحصائية.

تعريف: الفرض الإحصائي **Statistic Hypothesis**

هو ادعاء معيّن مبني على حيثيات معقولة حول معلمة من معالم المجتمع مثل المتوسط الحسابي μ أو الانحراف المعياري σ.

تعريف : المقياس الإحصائي

هو قيمة وحيدة محسوبة من العينة تحت شروط معينة.

تعريف: اختبارات الفروض الإحصائية (اختبار المعنوية)

هي طريقة معيارية لاختبار ادعاء ما حول معلمة من معالم المجتمع.

ملاحظة: سنكتفي في هذا الموضوع بدراسة معلمة واحدة من معالم المجتمع وهي المتوسط الحسابي μ.

إليك بعض الأمثلة عن الفروض التي يمكن اختبارها من خلال الطرق التي سنطوّرها في هذا الدرس.

على سبيل المثال:

- **في إدارة الأعمال:** تدّعي إحدى الصحف في مقال لها أنّ معظم الموظّفين يجدون عملًا عن طريق وكالات التوظيف.
- **في الطب:** يدّعي باحثون في الطبّ أنّ متوسّط درجة حرارة جسم أي بالغ معافى ليست ٣٧° سيليزية.
- **في سلامة الطيران المدنيّ:** تدّعي إدارة الطيران المدني في الكويت أن متوسط وزن المسافر (مع حقائبه) يتعدّى الوزن المسموح منذ عشرين سنة والبالغ ٨٤ كجم.

فرض العدم والفرض البديل Null and Alternative Hypotheses

- فرض العدم (ف٠): يفيد بأنّ قيمة معلمة المجتمع (مثل المتوسّط الحسابي μ) تساوي قيمة مزعومة. نختبر فرض العدم مباشرة أي نفترض بأنّه صحيح ونتوصّل إلى خلاصة برفض أو عدم رفض ف٠.
- الفرض البديل (ف١): يفيد بأن للمعلمة قيمة تختلف نوعًا ما عن فرض العدم (ف٠).
- يضم الشكل الرمزي للفرض البديل أحد هذه الرموز: $<$ أو $>$ أو $\neq$.

وستقتصر دراستنا على الحالة ($\neq$). فمثلًا: ف٠ : $\mu = ٩٨,٦$ ، ف١ : $\mu \neq ٩٨,٦$

الخطوات المتبعة لإجراء اختبار الفروض الإحصائية:

١ صياغة الفروض الإحصائية (فرض العدم ف٠ والفرض البديل ف١).

٢ التحقق من الانحراف المعياري للمجتمع σ (معلوم أم غير معلوم) وتحديد حجم العينة (ن) ومن ثم إيجاد المقياس الإحصائي للاختبار (ى أو ت)، (مسترشدًا بالجدول التالي):

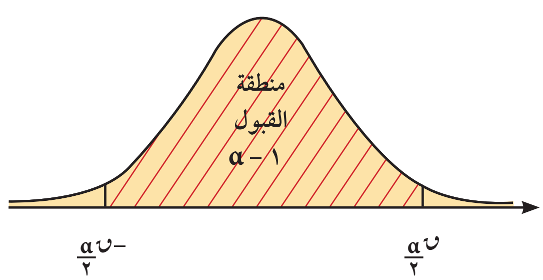
الانحراف المعياري (σ)	المقياس الإحصائي (ى أو ت)	حجم العينة (ن)
معلوم	$ى = \dfrac{\overline{س} - \mu}{\frac{\sigma}{\sqrt{ن}}}$	لا يشترط حجم معين للعينة
غير معلوم (نستبدل σ بـ ع)	$ى = \dfrac{\overline{س} - \mu}{\frac{ع}{\sqrt{ن}}}$	ن $>$ ٣٠
	$ت = \dfrac{\overline{س} - \mu}{\frac{ع}{\sqrt{ن}}}$	ن $\leq$ ٣٠

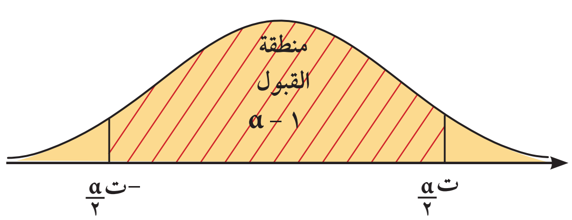
٣ تحديد مستوى المعنوية α وحساب القيمة الجدولية $ى_{\frac{\alpha}{٢}}$ من جدول التوزيع الطبيعي المعياري أو القيمة الجدولية $ت_{\frac{\alpha}{٢}}$ من جدول ت ذي درجات حرية.

٤ تحديد منطقة القبول: $\left(-ى_{\frac{\alpha}{٢}} ، ى_{\frac{\alpha}{٢}}\right)$ أو $\left(-ت_{\frac{\alpha}{٢}} ، ت_{\frac{\alpha}{٢}}\right)$ كما هو موضّح بالشكل.

٥ اتخاذ القرار الإحصائي (قبول فرض العدم) أو (رفض فرض العدم وقبول الفرض البديل).

ملاحظة: ستقتصر دراستنا على مستوى ثقة ٩٥٪





(١–٢–أ) إذا كان الانحراف المعياري لمجتمع σ معلوم

مثال (١)

تزعم شركة أن متوسط رواتب موظفيها يساوي ٤٠٠٠ دينار كويتي. إذا أخذت عينة من ٢٥ موظفًا، ووجد أن متوسط رواتب العينة هو ٣٩٥٠ دينارًا كويتيًّا فإذا علمت أن الانحراف المعياري للمجتمع σ = ١٢٥ دينارًا، وضّح كيفية إجراء الاختبار الإحصائي بمستوى ثقة ٩٥٪ .

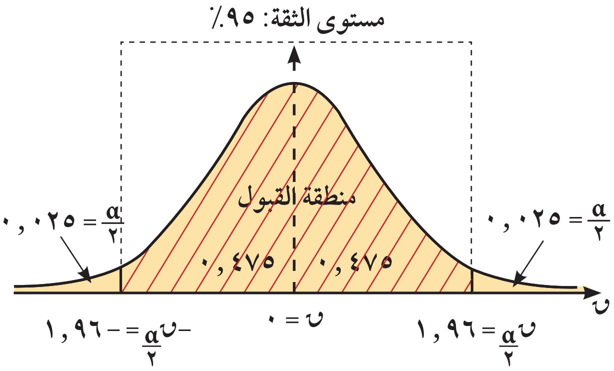
الحل:

① صياغة الفروض

ف٠ : μ = ٤٠٠٠ مقابل ف١ : $\mu \neq$ ٤٠٠٠

② ∵ σ = ١٢٥ (معلومة)

∴ نستخدم المقياس الإحصائي ى : $ى = \dfrac{\bar{س} - \mu}{\dfrac{\sigma}{\sqrt{ن}}}$



∵ ن = ٢٥ ، $\bar{س}$ = ٣٩٥٠

$ى = \dfrac{\bar{س} - \mu}{\dfrac{\sigma}{\sqrt{ن}}}$

∴ $ى = \dfrac{٣٩٥٠ - ٤٠٠٠}{\dfrac{١٢٥}{\sqrt{٢٥}}} = -٢$

③ ∵ مستوى الثقة ٩٥٪

∴ α = ٠,٠٥ ⟵ $\frac{\alpha}{٢}$ = ٠,٠٢٥

∴ $ى_{\frac{\alpha}{٢}}$ = ١,٩٦

④ منطقة القبول هي (−١,٩٦ ، ١,٩٦)

⑤ ∵ −٢ ∉ (−١,٩٦ ، ١,٩٦)

∴ القرار: نرفض فرض العدم μ = ٤٠٠٠ ونقبل الفرض البديل $\mu \neq$ ٤٠٠٠

حاول أن تحل

① يزعم صانع إطارات أن متوسط عمر الإطارات التي يصنعها μ = ٢٥٠٠٠ كم.

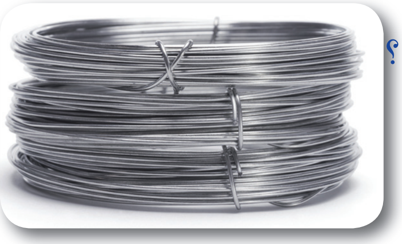
إذا أخذت عينة عشوائية من ١٥ إطارًا وأظهرت أن متوسطها الحسابي $\bar{س}$ = ٢٧٠٠٠ كم.

إذا علمت أن الانحراف المعياري للمجتمع = ٥٠٠٠ كم

فوضّح كيفية إجراء الاختبار الإحصائي لمستوى ثقة ٩٥٪ .

مثال (٢)

بيّنت الدراسة أن قوة تحمل أسلاك معدنية لها متوسط حسابي μ = ١٨٠٠ كجم مع انحراف معياري σ = ١٥٠ كجم. ويؤكد الأخصائيون في المصنع المنتج لهذه الاسلاك أن بإمكانهم زيادة قوة تحمل هذه الأسلاك، وتأكيدًا على ذلك تمّ اختبار عينة من ٤٠ سلكًا فتبيّن أن متوسط تحمل هذه الأسلاك يساوي ١٨٤٠ كجم.



هل يمكن قبول مثل هذا الفرض بمستوى معنوية α = ٠,٠٥؟

الحل:

١ صياغة الفروض:

ف. : μ = ١٨٠٠ مقابل ف١ : $\mu \neq$ ١٨٠٠

٢ ∵ σ = ١٥٠ (معلومة)

∴ نستخدم المقياس الإحصائي ى : $ى = \dfrac{\overline{س} - \mu}{\frac{\sigma}{\sqrt{ن}}}$

∵ ن = ٤٠ ، $\overline{س}$ = ١٨٤٠

$$ى = \dfrac{\overline{س} - \mu}{\frac{\sigma}{\sqrt{ن}}}$$

∴ $ى = \dfrac{١٨٤٠ - ١٨٠٠}{\frac{١٥٠}{\sqrt{٤٠}}} \approx ١,٦٨٦٥$

٣ ∴ α = ٠,٠٥ ⟵ $\frac{\alpha}{٢}$ = ٠,٠٢٥

∴ $ى_{\frac{\alpha}{٢}}$ = ١,٩٦

٤ منطقة القبول هي (-١,٩٦، ١,٩٦)

٥ ∵ ١,٦٨٦٥ ∈ (-١,٩٦، ١,٩٦)

∴ القرار بقبول فرض العدم μ = ١٨٠٠

حاول أن تحل

٢ متوسط العمر لعينة من ١٥٠ مصباحًا كهربائيًّا مصنعة في أحد المصانع هو $\overline{س}$ = ١٥٨٠ ساعة بانحراف معياري σ = ١٢٥ ساعة. يقول صاحب المصنع أن متوسط العمر μ = ١٦٢٠ ساعة.

اختبر الفرض μ = ١٦٢٠ ساعة مقابل الفرض $\mu \neq$ ١٦٢٠ ساعة باختيار مستوى معنوية α = ٠,٠٥

(١–٢–ب) إذا كان الانحراف المعياري لمجتمع σ غير معلوم، ن > ٣٠

مثال (٣)

إذا كانت ن = ٨٠ ، $\overline{\text{س}}$ = ٣٧,٢ ، عِ = ١,٧٩

اختبر الفرض بأن μ = ٣٧ عند مستوى معنوية α = ٠,٠٥

الحل:

١ صياغة الفروض

ف. : μ = ٣٧ مقابل ف١ : $\mu \neq$ ٣٧

٢ ∵ σ غير معلومة ، ن > ٣٠

∴ نستخدم المقياس الإحصائي ى : ى = $\dfrac{\overline{\text{س}} - \mu}{\frac{\text{عِ}}{\sqrt{\text{ن}}}}$

∵ ن = ٨٠ ، $\overline{\text{س}}$ = ٣٧,٢ ، عِ = ١,٧٩

∴ ى = $\dfrac{\overline{\text{س}} - \mu}{\frac{\text{عِ}}{\sqrt{\text{ن}}}}$

ى = $\dfrac{٣٧,٢ - ٣٧}{\frac{١,٧٩}{\sqrt{٨٠}}}$ = ٠,٩٩٩

٣ ∵ α = ٠,٠٥ ⟵ $\dfrac{\alpha}{٢}$ = ٠,٠٢٥

∴ $\text{ى}_{\frac{\alpha}{٢}}$ = ١,٩٦

٤ منطقة القبول هي (–١,٩٦ ، ١,٩٦)

٥ ∵ ٠,٩٩٩ ∈ (–١,٩٦ ، ١,٩٦)

∴ القرار بقبول فرض العدم μ = ٣٧

حاول أن تحل

٣ متوسط العمر لعينة من ١٠٠ مصباح كهربائي مصنعة في أحد المصانع $\overline{\text{س}}$ = ١ ٥٧٠ ساعة بانحراف معياري عِ = ١٢٠ ساعة. يقول صاحب المصنع إنّ متوسط العمر μ = ١ ٦٠٠ ساعة للمصابيح المصنعة في المصنع.

اختبر صحة الفرض μ = ١ ٦٠٠ ساعة مقابل الفرض $\mu \neq$ ١ ٦٠٠ ساعة وباختيار مستوى معنوية α = ٠,٠٥

(إرشاد: ف. : μ = ١ ٦٠٠ ، ف١ : $\mu \neq$ ١ ٦٠٠).

(١–٢–ج) إذا كان الانحراف المعياري للمجتمع σ غير معلوم ، ن ≥ ٣٠

مثال (٤)

يعتقد مدير شركة دراسات إحصائية أن متوسط الإنفاق الشهري على الطعام في منازل مدينة معينة يساوي ٢٩٠ دينارًا كويتيًّا.
فإذا أخذت عينة عشوائية من ١٠ منازل تبيّن أن متوسطها الحسابي $\overline{س}$ = ٢٨٣ دينارًا وانحرافها المعياري $ع_{ر}$ = ٣٢ دينارًا.
فهل يمكن الاعتماد على هذه العينة لتأكيد ما افترضه؟
استخدم مستوى ثقة ٩٥ ٪ (علمًا بأن المجتمع يتبع توزيعًا طبيعيًّا).

الحل:

١ صياغة الفروض: $ف_{٠}$: μ = ٢٩٠ مقابل $ف_{١}$: $\mu \neq$ ٢٩٠

٢ ∵ σ غير معلومة ، ن = ١٠ (ن ≥ ٣٠)

∴ نستخدم المقياس الإحصائي ت : $ت = \dfrac{\overline{س} - \mu}{\dfrac{ع_{ر}}{\sqrt{ن}}}$

∵ ن = ١٠ ، $\overline{س}$ = ٢٨٣ ، $ع_{ر}$ = ٣٢

$$ت = \dfrac{\overline{س} - \mu}{\dfrac{ع_{ر}}{\sqrt{ن}}}$$

$$\therefore ت = \dfrac{٢٨٣ - ٢٩٠}{\dfrac{٣٢}{\sqrt{١٠}}} = -٠{,}٦٩١٧$$

٣ ∵ مستوى الثقة ٩٥٪ ، درجات الحرية (ن – ١) = ١٠ – ١ = ٩

∴ α = ٠,٠٥ ⟵ $\dfrac{\alpha}{٢}$ = ٠,٠٢٥

∴ $ت_{\frac{\alpha}{٢}}$ = ٢,٢٦٢

٤ منطقة القبول هي (–٢,٢٦٢ ، ٢,٢٦٢)

٥ ∵ – ٠,٦٩١٧ ∈ (–٢,٢٦٢ ، ٢,٢٦٢)

∴ القرار بقبول فرض العدم μ = ٢٩٠

∴ يمكن الاعتماد على هذه العينة.

حاول أن تحل

٤ في المثال (٤)، إذا أجريت دراسة إحصائية أخرى على المدينة ذاتها وتبيّن من خلالها أن $\overline{س}$ = ٢٩٦ ، $ع_{ر}$ = ٥ لعينة من ١٠ منازل مع استخدام درجة الثقة نفسها.
فهل يبقى افتراض المدير عند الشركة صحيحًا أم لا؟ وضّح إجابتك.

المرشد لحل المسائل

نظرًا لأهميّة المياه بالنسبة إلى صحة الإنسان وحياته، قررت مؤسسة تعنى بذلك، القيام بحملة تهدف إلى التأكد من أن كل شخص يستهلك متوسط قدره ٢٠٠٠ ملل يوميًّا من مياه الشرب.

في دراسة سابقة لعينة من ١٠٠ شخص، لاحظت المؤسسة أنّ المتوسط الحسابي للاستهلاك: $\overline{س} = ١٨٥٠$ ملل مع انحراف معياري ع = ٩٠٠ ملل.

وفي دراسة جديدة لعينة من ١٠٠ شخص، وبعد القيام بحملتها، لاحظت أن المتوسط الحسابي للاستهلاك: $\overline{س} = ١٩٠٠$ ملل مع انحراف معياري ع = ٣٠٠ ملل.

اعتقدت المؤسسة أنّ حملتها قد نجحت بما أنّ المتوسط الحسابي للاستهلاك قد ازداد ٥٠ ملل وقد اقترب كثيرًا من هدفها وهو ٢٠٠٠ ملل يوميًّا للشخص الواحد. هل المؤسسة على حق؟ اشرح.

الحل:

وضع يوسف جدولًا ليختبر فرضية الشركة من خلال اختبارات إحصائية مع:

$ف_{٠} : \mu = ٢٠٠٠$ مقابل $ف_{١} : \mu \neq ٢٠٠٠$، ومستوى الثقة ٠,٩٥

	الدراسة السابقة	الدراسة الجديدة
المعايير	$\overline{س} = ١٨٥٠$ ، ع = ٩٠٠ ، ن = ١٠٠	$\overline{س} = ١٩٠٠$ ، ع = ٣٠٠ ، ن = ١٠٠
القيمة الجدولية	$ى_{\frac{\alpha}{٢}} = ١,٩٦$	$ى_{\frac{\alpha}{٢}} = ١,٩٦$
قيمة الاختبار الإحصائي	$ى = \frac{\overline{س} - \mu}{\frac{ع}{\sqrt{ن}}} = \frac{١٨٥٠ - ٢٠٠٠}{\frac{٩٠٠}{\sqrt{١٠٠}}} = -١,٦٦$	$ى = \frac{١٩٠٠ - ٢٠٠٠}{\frac{٣٠٠}{\sqrt{١٠٠}}} = -٣,٣٣$
الفترة	(−١,٩٦، ١,٩٦)	(−١,٩٦، ١,٩٦)
القرار	$\because -١,٦٦ \in (-١,٩٦، ١,٩٦)$ قبول $ف_{٠} : \mu = ٢٠٠٠$ ملل يوميًّا	$\because -٣,٣٣ \notin (-١,٩٦٦، ١,٩٦٦)$ رفض $ف_{٠}$ والأخذ بـ $ف_{١}$: $\mu \neq ٢٠٠٠$ ملل يوميًّا

الاستنتاج:

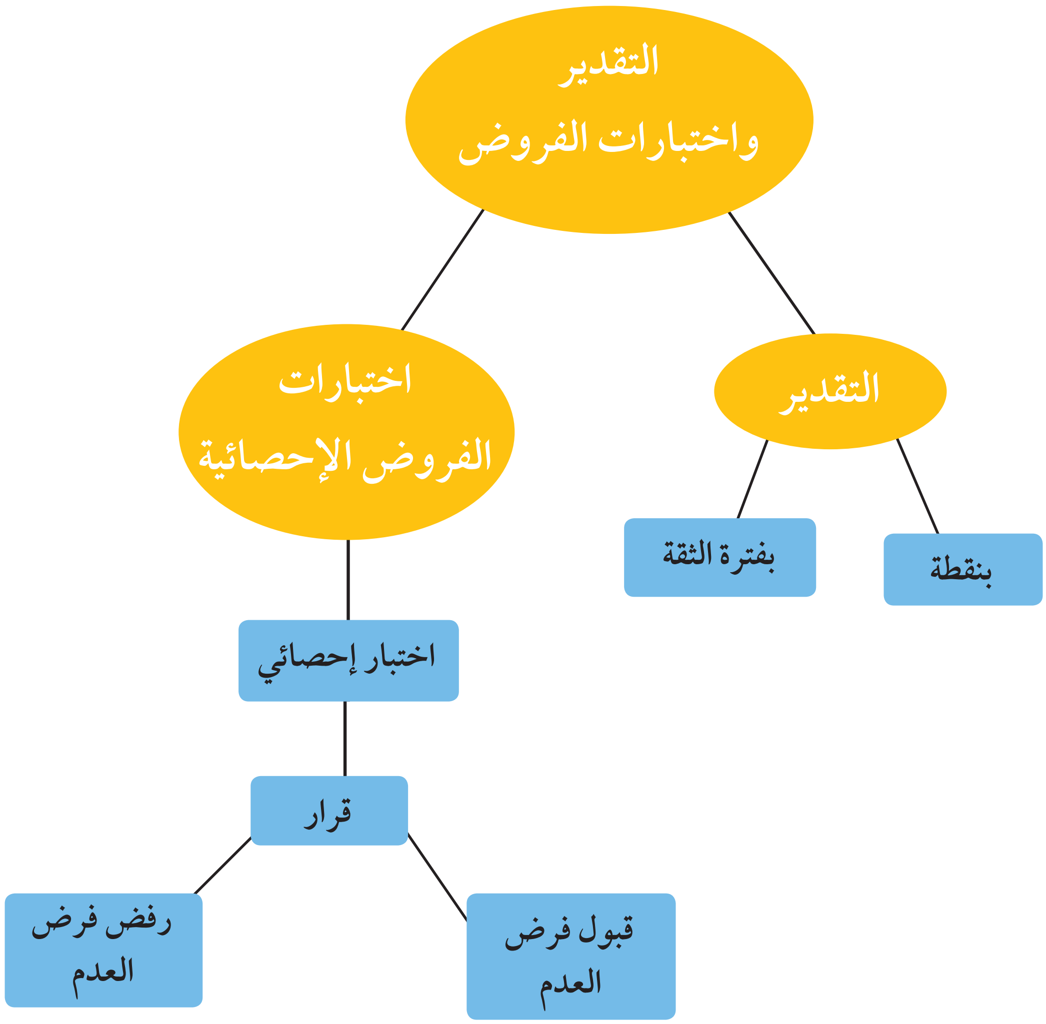
لم تكن الحملة ضرورية، والحصول على قيمة متوسطة أكبر لا يعني الاقتراب من الهدف المنشود.

مسألة إضافية

قامت مؤسسة أخرى بحملة على عينة من ١٠٠ شخص تهدف إلى التأكد من أنّ المتوسط الحسابي لاستهلاك كل شخص لمياه الشرب $\mu = ٢٠٠٠$ ملل يوميًّا. فأتت النتائج على الشكل التالي:

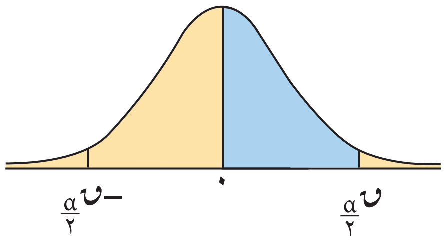
$\overline{س} = ٢١٠٠$ ملل، ع = ٨٠٠ ملل. برأيك، هل كانت حملة هذه المؤسسة ناجحة؟

مخطط تنظيمي للوحدة الأولى



ملخص

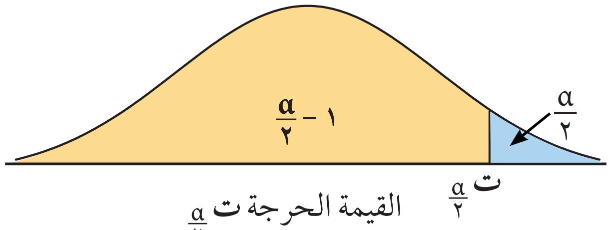
- المعلمة هي ثابت يصف المجتمع أو يصف توزيع المجتمع كالمتوسط الحسابي μ أو الانحراف المعياري σ.
- الإحصاءة هو اقتران تتعيّن قيمته من العينة كالمتوسط الحسابي $\overline{س}$ أو الانحراف المعياري ع.
- التقدير بنقطة هي قيمة وحيدة محسوبة من العينة تستخدم لتقدير معلمة مجهولة من معالم المجتمع.
- فترة الثقة هي فترة طرفاها متغيران عشوائيان (أي أنها فترة عشوائية) تحوي إحدى معالم المجتمع بنسبة معينة تسمى درجة الثقة (مستوى الثقة).
- α هي درجة (نسبة) الخطأ في التقدير.
- مستوى الثقة ١٠٠٪ ($١ - \alpha$) ويسمى ($١ - \alpha$) معامل مستوى الثقة.
- $ن_{\frac{\alpha}{٢}}$ هي القيمة الحرجة المستخرجة من جدول التوزيع الطبيعي المعياري.
- $\overline{س}$ هو المتوسط الحسابي للعينة.
- ع هو الانحراف المعياري للعينة.
- $ت_{\frac{\alpha}{٢}}$ هي القيمة الحرجة المستخرجة من جدول التوزيع ت.
- هامش الخطأ هـ = $ن_{\frac{\alpha}{٢}} \cdot \frac{\sigma}{\sqrt{ن}}$ في حالة الانحراف المعياري σ معلوم والتوزيع الطبيعي.
- هامش الخطأ هـ = $ن_{\frac{\alpha}{٢}} \cdot \frac{ع}{\sqrt{ن}}$ في حالة الانحراف المعياري σ غير معلوم ون > ٣٠ والتوزيع الطبيعي.
- هامش الخطأ هـ = $ت_{\frac{\alpha}{٢}} \cdot \frac{ع}{\sqrt{ن}}$، إذا كانت σ غير معلوم و ن ≤ ٣٠ والتوزيع طبيعي.
- الفرض الإحصائي: هو ادعاء معين مبني على حيثيات معقولة حول معلمة من معالم المجتمع مثل المتوسط الحسابي μ أو الانحراف المعياري σ.
- المقياس الإحصائي هو قيمة وحيدة محسوبة من العينة تحت شروط معينة.
- اختبارات الفروض الإحصائية (اختبار المعنوية) هي طريقة معيارية لاختبار ادعاء ما حول معلمة من معالم المجتمع.



جدول التوزيع الطبيعي المعياري (ن)

ن	٠,٠٠	٠,٠١	٠,٠٢	٠,٠٣	٠,٠٤	٠,٠٥	٠,٠٦	٠,٠٧	٠,٠٨	٠,٠٩
٠,٠	٠,٠٠٠٠	٠,٠٠٤٠	٠,٠٠٨٠	٠,٠١٢٠	٠,٠١٦٠	٠,٠١٩٩	٠,٠٢٣٩	٠,٠٢٧٩	٠,٠٣١٩	٠,٠٣٥٩
٠,١	٠,٠٣٩٨	٠,٠٤٣٨	٠,٠٤٧٨	٠,٠٥١٧	٠,٠٥٥٧	٠,٠٥٩٦	٠,٠٦٣٦	٠,٠٦٧٥	٠,٠٧١٤	٠,٠٧٥٣
٠,٢	٠,٠٧٩٣	٠,٠٨٣٢	٠,٠٨٧١	٠,٠٩١٠	٠,٠٩٤٨	٠,٠٩٨٧	٠,١٠٢٦	٠,١٠٦٤	٠,١١٠٣	٠,١١٤١
٠,٣	٠,١١٧٩	٠,١٢١٧	٠,١٢٥٥	٠,١٢٩٣	٠,١٣٣١	٠,١٣٦٨	٠,١٤٠٦	٠,١٤٤٣	٠,١٤٨٠	٠,١٥١٧
٠,٤	٠,١٥٥٤	٠,١٥٩١	٠,١٦٢٨	٠,١٦٦٤	٠,١٧٠٠	٠,١٧٣٦	٠,١٧٧٢	٠,١٨٠٨	٠,١٨٤٤	٠,١٨٧٩
٠,٥	٠,١٩١٥	٠,١٩٥٠	٠,١٩٨٥	٠,٢٠١٩	٠,٢٠٥٤	٠,٢٠٨٨	٠,٢١٢٣	٠,٢١٥٧	٠,٢١٩٠	٠,٢٢٢٤
٠,٦	٠,٢٢٥٧	٠,٢٢٩١	٠,٢٣٢٤	٠,٢٣٥٧	٠,٢٣٨٩	٠,٢٤٢٢	٠,٢٤٥٤	٠,٢٤٨٦	٠,٢٥١٧	٠,٢٥٤٩
٠,٧	٠,٢٥٨٠	٠,٢٦١١	٠,٢٦٤٢	٠,٢٦٧٣	٠,٢٧٠٤	٠,٢٧٣٤	٠,٢٧٦٤	٠,٢٧٩٤	٠,٢٨٢٣	٠,٢٨٥٢
٠,٨	٠,٢٨٨١	٠,٢٩١٠	٠,٢٩٣٩	٠,٢٩٦٧	٠,٢٩٩٥	٠,٣٠٢٣	٠,٣٠٥١	٠,٣٠٧٨	٠,٣١٠٦	٠,٣١٣٣
٠,٩	٠,٣١٥٩	٠,٣١٨٦	٠,٣٢١٢	٠,٣٢٣٨	٠,٣٢٦٤	٠,٣٢٨٩	٠,٣٣١٥	٠,٣٣٤٠	٠,٣٣٦٥	٠,٣٣٨٩
١,٠	٠,٣٤١٣	٠,٣٤٣٨	٠,٣٤٦١	٠,٣٤٨٥	٠,٣٥٠٨	٠,٣٥٣١	٠,٣٥٥٤	٠,٣٥٧٧	٠,٣٥٩٩	٠,٣٦٢١
١,١	٠,٣٦٤٣	٠,٣٦٦٥	٠,٣٦٨٦	٠,٣٧٠٨	٠,٣٧٢٩	٠,٣٧٤٩	٠,٣٧٧٠	٠,٣٧٩٠	٠,٣٨١٠	٠,٣٨٣٠
١,٢	٠,٣٨٤٩	٠,٣٨٦٩	٠,٣٨٨٨	٠,٣٩٠٧	٠,٣٩٢٥	٠,٣٩٤٤	٠,٣٩٦٢	٠,٣٩٨٠	٠,٣٩٩٧	٠,٤٠١٥
١,٣	٠,٤٠٣٢	٠,٤٠٤٩	٠,٤٠٦٦	٠,٤٠٨٢	٠,٤٠٩٩	٠,٤١١٥	٠,٤١٣١	٠,٤١٤٧	٠,٤١٦٢	٠,٤١٧٧
١,٤	٠,٤١٩٢	٠,٤٢٠٧	٠,٤٢٢٢	٠,٤٢٣٦	٠,٤٢٥١	٠,٤٢٦٥	٠,٤٢٧٩	٠,٤٢٩٢	٠,٤٣٠٦	٠,٤٣١٩
١,٥	٠,٤٣٣٢	٠,٤٣٤٥	٠,٤٣٥٧	٠,٤٣٧٠	٠,٤٣٨٢	٠,٤٣٩٤	٠,٤٤٠٦	٠,٤٤١٨	٠,٤٤٢٩	٠,٤٤٤١
١,٦	٠,٤٤٥٢	٠,٤٤٦٣	٠,٤٤٧٤	٠,٤٤٨٤	٠,٤٤٩٥	٠,٤٥٠٥	٠,٤٥١٥	٠,٤٥٢٥	٠,٤٥٣٥	٠,٤٥٤٥
١,٧	٠,٤٥٥٤	٠,٤٥٦٤	٠,٤٥٧٣	٠,٤٥٨٢	٠,٤٥٩١	٠,٤٥٩٩	٠,٤٦٠٨	٠,٤٦١٦	٠,٤٦٢٥	٠,٤٦٣٣
١,٨	٠,٤٦٤١	٠,٤٦٤٩	٠,٤٦٥٦	٠,٤٦٦٤	٠,٤٦٧١	٠,٤٦٧٨	٠,٤٦٨٦	٠,٤٦٩٣	٠,٤٦٩٩	٠,٤٧٠٦
١,٩	٠,٤٧١٣	٠,٤٧١٩	٠,٤٧٢٦	٠,٤٧٣٢	٠,٤٧٣٨	٠,٤٧٤٤	٠,٤٧٥٠	٠,٤٧٥٦	٠,٤٧٦١	٠,٤٧٦٧
٢,٠	٠,٤٧٧٢	٠,٤٧٧٨	٠,٤٧٨٣	٠,٤٧٨٨	٠,٤٧٩٣	٠,٤٧٩٨	٠,٤٨٠٣	٠,٤٨٠٨	٠,٤٨١٢	٠,٤٨١٧
٢,١	٠,٤٨٢١	٠,٤٨٢٦	٠,٤٨٣٠	٠,٤٨٣٤	٠,٤٨٣٨	٠,٤٨٤٢	٠,٤٨٤٦	٠,٤٨٥٠	٠,٤٨٥٤	٠,٤٨٥٧
٢,٢	٠,٤٨٦١	٠,٤٨٦٤	٠,٤٨٦٨	٠,٤٨٧١	٠,٤٨٧٥	٠,٤٨٧٨	٠,٤٨٨١	٠,٤٨٨٤	٠,٤٨٨٧	٠,٤٨٩٠
٢,٣	٠,٤٨٩٣	٠,٤٨٩٦	٠,٤٨٩٨	٠,٤٩٠١	٠,٤٩٠٤	٠,٤٩٠٦	٠,٤٩٠٩	٠,٤٩١١	٠,٤٩١٣	٠,٤٩١٦
٢,٤	٠,٤٩١٨	٠,٤٩٢٠	٠,٤٩٢٢	٠,٤٩٢٥	٠,٤٩٢٧	٠,٤٩٢٩	٠,٤٩٣١	٠,٤٩٣٢	٠,٤٩٣٤	٠,٤٩٣٦
٢,٥	٠,٤٩٣٨	٠,٤٩٤٠	٠,٤٩٤١	٠,٤٩٤٣	٠,٤٩٤٥	٠,٤٩٤٦	٠,٤٩٤٨	٠,٤٩٤٩	٠,٤٩٥١	٠,٤٩٥٢
٢,٦	٠,٤٩٥٣	٠,٤٩٥٥	٠,٤٩٥٦	٠,٤٩٥٧	٠,٤٩٥٩	٠,٤٩٦٠	٠,٤٩٦١	٠,٤٩٦٢	٠,٤٩٦٣	٠,٤٩٦٤
٢,٧	٠,٤٩٦٥	٠,٤٩٦٦	٠,٤٩٦٧	٠,٤٩٦٨	٠,٤٩٦٩	٠,٤٩٧٠	٠,٤٩٧١	٠,٤٩٧٢	٠,٤٩٧٣	٠,٤٩٧٤
٢,٨	٠,٤٩٧٤	٠,٤٩٧٥	٠,٤٩٧٦	٠,٤٩٧٧	٠,٤٩٧٧	٠,٤٩٧٨	٠,٤٩٧٩	٠,٤٩٧٩	٠,٤٩٨٠	٠,٤٩٨١
٢,٩	٠,٤٩٨١	٠,٤٩٨٢	٠,٤٩٨٢	٠,٤٩٨٣	٠,٤٩٨٤	٠,٤٩٨٤	٠,٤٩٨٥	٠,٤٩٨٥	٠,٤٩٨٦	٠,٤٩٨٦
٣,٠	٠,٤٩٨٧	٠,٤٩٨٧	٠,٤٩٨٧	٠,٤٩٨٨	٠,٤٩٨٨	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٨٩	٠,٤٩٩٠	٠,٤٩٩٠
٣,١٠ وأكثر	٠,٤٩٩٩									

ملاحظة: استخدم ٠,٤٩٩٩ عندما تزيد قيمة ن عن ٣,٠٩



القيمة الحرجة ت $\frac{\alpha}{٢}$

جدول التوزيع ت						
	$\frac{\alpha}{٢}$					
درجات الحرية (ن – ١)	٠,٠٠٥	٠,٠١	٠,٠٢٥	٠,٠٥	٠,١٠	٠,٢٥
١	٦٣,٦٥٧	٣١,٨٢١	١٢,٧٠٦	٦,٣١٤	٣,٠٧٨	١,٠٠٠
٢	٩,٩٢٥	٦,٩٦٥	٤,٣٠٣	٢,٩٢٠	١,٨٨٦	٠,٨١٦
٣	٥,٨٤١	٤,٥٤١	٣,١٨٢	٢,٣٥٣	١,٦٣٨	٠,٧٦٥
٤	٤,٦٠٤	٣,٧٤٧	٢,٧٧٦	٢,١٣٢	١,٥٣٣	٠,٧٤١
٥	٤,٠٣٢	٣,٣٦٥	٢,٥٧١	٢,٠١٥	١,٤٧٦	٠,٧٢٧
٦	٣,٧٠٧	٣,١٤٣	٢,٤٤٧	١,٩٤٣	١,٤٤٠	٠,٧١٨
٧	٣,٥٠٠	٢,٩٩٨	٢,٣٦٥	١,٨٩٥	١,٤١٥	٠,٧١١
٨	٣,٣٥٥	٢,٨٩٦	٢,٣٠٦	١,٨٦٠	١,٣٩٧	٠,٧٠٦
٩	٣,٢٥٠	٢,٨٢١	٢,٢٦٢	١,٨٣٣	١,٣٨٣	٠,٧٠٣
١٠	٣,١٦٩	٢,٧٦٤	٢,٢٢٨	١,٨١٢	١,٣٧٢	٠,٧٠٠
١١	٣,١٠٦	٢,٧١٨	٢,٢٠١	١,٧٩٦	١,٣٦٣	٠,٦٩٧
١٢	٣,٠٥٤	٢,٦٨١	٢,١٧٩	١,٧٨٢	١,٣٥٦	٠,٦٩٦
١٣	٣,٠١٢	٢,٦٥٠	٢,١٦٠	١,٧٧١	١,٣٥٠	٠,٦٩٤
١٤	٢,٩٧٧	٢,٦٢٥	٢,١٤٥	١,٧٦١	١,٣٤٥	٠,٦٩٢
١٥	٢,٩٤٧	٢,٦٠٢	٢,١٣٢	١,٧٥٣	١,٣٤١	٠,٦٩١
١٦	٢,٩٢١	٢,٥٨٤	٢,١٢٠	١,٧٤٦	١,٣٣٧	٠,٦٩٠
١٧	٢,٨٩٨	٢,٥٦٧	٢,١١٠	١,٧٤٠	١,٣٣٣	٠,٦٨٩
١٨	٢,٨٧٨	٢,٥٥٢	٢,١٠١	١,٧٣٤	١,٣٣٠	٠,٦٨٨
١٩	٢,٨٦١	٢,٥٤٠	٢,٠٩٣	١,٧٢٩	١,٣٢٨	٠,٦٨٨
٢٠	٢,٨٤٥	٢,٥٢٨	٢,٠٨٦	١,٧٢٥	١,٣٢٥	٠,٦٨٧
٢١	٢,٨٣١	٢,٥١٨	٢,٠٨٠	١,٧٢١	١,٣٢٣	٠,٦٨٦
٢٢	٢,٨١٩	٢,٥٠٨	٢,٠٧٤	١,٧١٧	١,٣٢١	٠,٦٨٦
٢٣	٢,٨٠٧	٢,٥٠٠	٢,٠٦٩	١,٧١٤	١,٣٢٠	٠,٦٨٥
٢٤	٢,٧٩٧	٢,٤٩٢	٢,٠٦٤	١,٧١١	١,٣١٨	٠,٦٨٥
٢٥	٢,٧٨٧	٢,٤٨٥	٢,٠٦٠	١,٧٠٨	١,٣١٦	٠,٦٨٤
٢٦	٢,٧٧٩	٢,٤٧٩	٢,٠٥٦	١,٧٠٦	١,٣١٥	٠,٦٨٤
٢٧	٢,٧٧١	٢,٤٧٣	٢,٠٥٢	١,٧٠٣	١,٣١٤	٠,٦٨٤
٢٨	٢,٧٦٣	٢,٤٦٧	٢,٠٤٨	١,٧٠١	١,٣١٣	٠,٦٨٣
٢٩	٢,٧٥٦	٢,٤٦٢	٢,٠٤٥	١,٦٩٩	١,٣١١	٠,٦٨٣
٣٠ وأكثر	٢,٥٧٥	٢,٣٢٧	١,٩٦٠	١,٦٤٥	١,٢٨٢	٠,٦٧٥

الوحدة
الثانية

الارتباط والانحدار

Correlation and Regression

مشروع الوحدة: ضغط الدم

١ مقدمة المشروع: يعتبر ضغط الدم عند الإنسان من أهم العوامل المؤثرة في حياة كل شخص. إن قياس ضغط الدم لجهة ارتفاعه أو انخفاضه عن معدله العام يساعد على المعالجة المبكرة وبالتالي التخفيف قدر الإمكان من حدوث النوبات القلبية المفاجئة. علمًا أن وزارة الصحة في دولة الكويت قد نبهت إلى عوارض ارتفاع ضغط الدم وخصوصًا لدى المسنين وأصحاب السمنة.

٢ الهدف: دراسة العلاقة بين وزن عدد من الأفراد (بالكيلوجرام) ومعدل ضغط الدم لديهم وذلك بتنفيذ ما يلي:

أ زيارة إحدى العيادات الطبية لتكوين جدول يبيّن وزن عدد من الأشخاص (ذكور) ومعدل ضغط الدم المقابل لكل وزن.

ب زيارة إحدى المستشفيات لتكوين جدول يبيّن وزن عدد من الأشخاص (إناث) ومعدل ضغط الدم المقابل لكل وزن.

٣ اللوازم: آلة حاسبة - ورق رسم بياني.

٤ أسئلة حول التطبيق:

أ كم عدد الأشخاص في العينة التي سوف تختارها في العيادة أو في المستشفى؟ احرص على أن يكون العدد نفسه في الحالتين.

ب مثّل على ورق رسم بياني مخطط انتشار لنتائج جدول العيادة وعلى ورق رسم بياني آخر مخطط انتشار لنتائج جدول المستشفى.

جـ هل يوجد لكل مخطط انتشار علاقة تصاعدية أو تنازلية بين الوزن ومعدل ضغط الدم؟ اشرح.

د من كل جدول لنأخذ ن (عدد الأشخاص)، س (الوزن)، ص (معدل ضغط الدم).

أوجد: س^2، مجـ س^2، ص^2، مجـ ص^2، س ص، مجـ س ص، مجـ س، مجـ ص.

هـ لكل جدول استنتج قيمة ما يلي:

$$\frac{\text{ن مجـ س ص} - \text{مجـ س مجـ ص}}{\sqrt{\text{ن مجـ س}^2 - (\text{مجـ س})^2}\ \sqrt{\text{ن مجـ ص}^2 - (\text{مجـ ص})^2}}$$

و ماذا تلاحظ لكل قيمة وجدتها؟ اشرح.

٥ التقرير: اكتب تقريرًا مفصّلًا يوضّح النتائج التي توصلت إليها عارضًا اقتراحاتك ونصائحك عن علاقة الوزن بمعدل ضغط الدم. هل ترى أي ترابط بين كل مخطط انتشار والقيمة المقابلة التي وجدتها؟

دروس الوحدة

٢-١ الارتباط	٢-٢الانحدار
(٢-١-أ) المخطط الانتشاري	
(٢-١-ب) معامل الارتباط الخطي	

الوحدة الثانية

أين أنت الآن (المعارف السابقة المكتسبة)

- التقدير بنقطة.
- التقدير بفترة ثقة.
- الفروض الإحصائية.
- الاختبارات الإحصائية.

ماذا سوف تتعلم؟

- الارتباط.
- مخطط الانتشار.
- مُعامل الارتباط (بيرسون).
- تحليل مُعامل الارتباط.
- الانحدار ومعادلته.
- توقع قيمة أحد المتغيرين.

المصطلحات الأساسية

الارتباط – مخطط الانتشار – مُعامل ارتباط بيرسون – ارتباط طردي (موجب) تام – ارتباط عكسي (سالب) تام – ارتباط منعدم – ارتباط طردي (موجب) قوي – ارتباط طردي (موجب) متوسط – ارتباط طردي (موجب) ضعيف – ارتباط عكسي (سالب) ضعيف – ارتباط عكسي (سالب) متوسط – ارتباط عكسي (سالب) قوي – الانحدار – معادلة خط الانحدار.

أضف إلى معلوماتك

يعتقد بعض الناس أنه بإمكانهم توقع طول العمر ومعرفته بالنظر إلى طول خط الحياة في كف يدهم. لكن إحدى الدراسات الطبّية أثبتت أنه لا وجود لرابط أو علاقة بين طول خط الحياة في كف الإنسان وطول عمره، وأن ما اعتقده ومازال يعتقده البعض عارٍ عن الصحة.

٢ ــ ١

الارتباط

Correlation

سوف تتعلم

- مفهوم الارتباط.
- رسم مخطط الانتشار.
- إيجاد مُعامل ارتباط بيرسون.
- تحليل قيمة مُعامل الارتباط.
- توقع قيمة أحد المتغيرين.

دعنا نفكر ونتناقش

هل تساءلت يومًا: كيف تحسب العلاقة بين الطول والوزن؟

ما الذي يربط بين التدخين والإصابة بمرض السرطان؟

كيف نجد رابطًا بين وزن سيارة واستهلاكها للوقود؟

كيف يتغيّر سعر الذهب مع تغير قيمة الدولار الأمريكي؟

وما هي أفضل وسيلة للتقدير لنقترب من الحقيقة؟

الارتباط Correlation

من دراستنا السابقة تمّ عرض بعض المقاييس الإحصائية، مثل: مقاييس النزعة المركزية (المتوسط الحسابي – الوسيط – المنوال) ومقاييس التشتت (المدى – التباين – الانحراف المعياري).

نلاحظ أن هذه المقاييس كانت تصف شكل البيانات التي تمّ جمعها من ظاهرة إحصائية واحدة أي من متغير واحد والذي يمكن الحصول عليه من العينة. بينما يقابلنا في حياتنا العملية مواقف كثيرة تتضمن متغيرين (ظاهرتين) أو أكثر ويكون تساؤلنا: هل هناك علاقة بين هذه المتغيرات؟ وما هو شكل هذه العلاقة؟ وأيضًا كيف يمكن التنبؤ بقيمة أحد هذين المتغيرين إذا علم قيمة المتغير الآخر؟ وكثيرًا ما يرى الباحثون ضرورة دراسة العلاقة بين متغيرين (ظاهرتين) كما يتضح من الأمثلة التالية:

- الطول والوزن.
- التدخين والإصابة بمرض السرطان.
- وزن سيارة واستهلاكها للوقود.
- الإنفاق والدخل.
- سعر السلعة والكمية المعروضة منها.
- العمر وضغط الدم.

والأمثلة في هذا المجال كثيرة ومتعددة، ولدراسة العلاقة بين هذه الظواهر ندرس ما يسمى الارتباط.

تعريف: الارتباط

هو العلاقة بين متغيرين.

سنرمز للمتغير الأول بالرمز «س»، وهو المتغير الذي يتم تحديده من قبل الباحث القائم بالدراسة ويسمى «بالمتغير المستقل».

ونرمز للمتغير الثاني بالرمز «ص»، وهذا المتغير غير مستقل بذاته لأن نتيجته مرتبطة بالمتغير المستقل ولذلك يسمى «بالمتغير التابع».

(٢–١–أ) المخطط الانتشاري Scatter Plot

تعريف: المخطط الانتشاري

هو عبارة عن تمثيل بياني لعدد من الأزواج المرتبة (س، ص) تستخدم لوصف العلاقة بين المتغيرين.

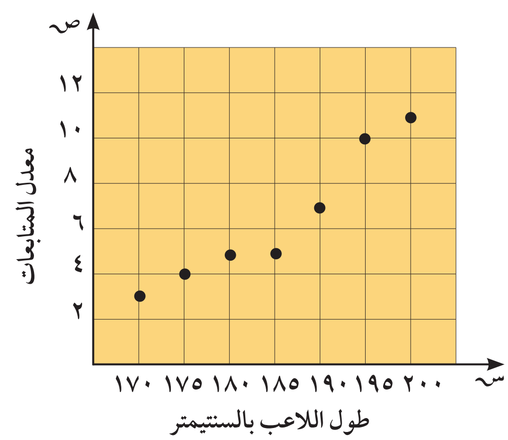
مثال (١)

الجدول التالي يوضح العلاقة بين طول اللاعب (س) ومعدّل المتابعات (ص)، لسبعة لاعبين في مباراة كرة السلة.

طول اللاعب (بالسنتيمتر) (س)	١٧٠	١٧٥	١٨٠	١٨٥	١٩٠	١٩٥	٢٠٠
معدل المتابعات (ص)	٣	٤	٥	٥	٧	١٠	١١

المطلوب: ارسم المخطط الانتشاري.

الحل:



حاول أن تحل

١ ارسم مخطط الانتشار الذي يوضّح البيانات التالية:

س	١٠٠	١١٠	١٢٠	١٣٠	١٤٠	١٦٠	١٧٠	١٨٠	١٩٠
ص	٢٢	٢٠	٢٠	١٩	١٨	١٧	١٥	١٦	١٤

أنواع الارتباط

١ ارتباط طردي (موجب):

هو علاقة بين متغيرين س، ص بحيث إذا تغير المتغير المستقل (س) فإن المتغير التابع (ص) يتبعه في نفس الاتجاه.

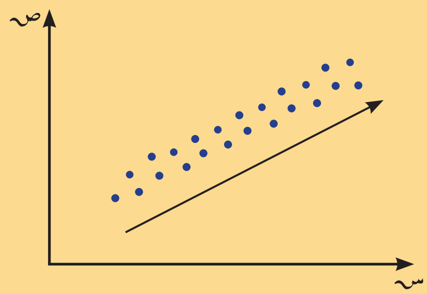
أي أنه كلما زادت قيمة س تزداد تبعًا لها قيمة ص.

٢ ارتباط عكسي (سالب):

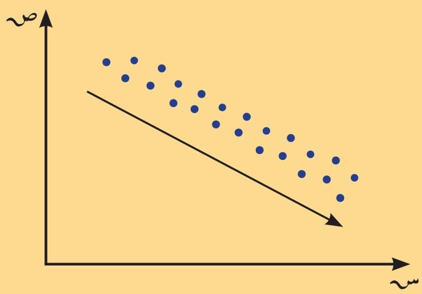
هو علاقة بين متغيرين س، ص بحيث إذا تغير المتغير المستقل (س) فإن المتغير التابع (ص) يتبعه في الاتجاه المضاد.

أي أنه كلما زادت قيمة س تتناقص تبعًا لها قيمة ص.

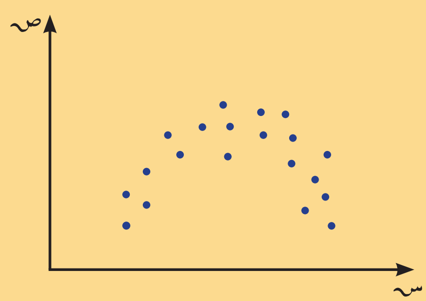
بعض الأشكال التي توضّح أنواع الارتباط



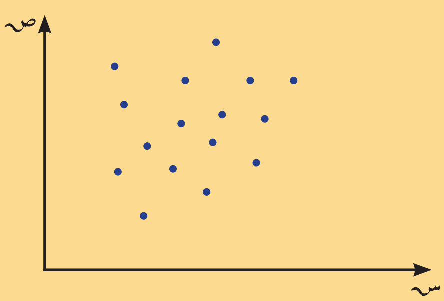
علاقة خطيّة
ارتباط طردي (موجب)



علاقة خطيّة
ارتباط عكسي (سالب)



علاقة غير خطيّة



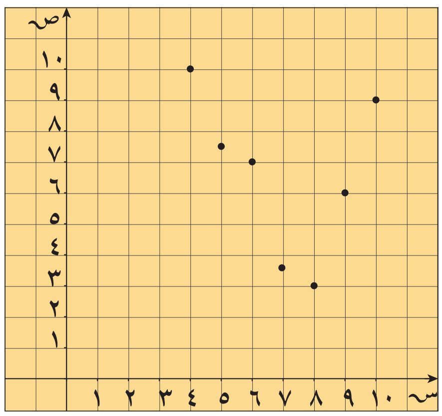
لا توجد علاقة

مثال (٢)

ارسم مخطط الانتشار للبيانات التالية وحدّد نوع العلاقة الّتي تعبّر عنها.

س	٤	٥	٦	٧	٨	٩	١٠
ص	١٠	٧,٥	٧	٣,٥	٣	٦	٩

الحل:



لا توجد علاقة.

حاول أن تحل

٢ ارسم مخطط الانتشار للبيانات التالية وحدّد نوع العلاقة التي تعبّر عنها.

س	١	٢	٣	٤	٥	٦	٧
ص	٢	٤	٦	٨	١٠	١٢	١٤

مثال (٣)

البيانات التالية تبيّن العلاقة بين عمر الشخص وعدد ساعات التمرينات الرياضية التي يقوم بها:

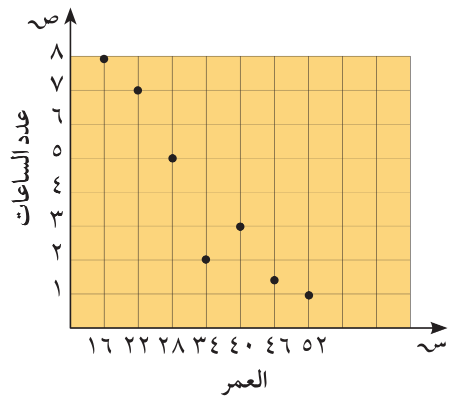
العمر (س)	١٦	٢٢	٢٨	٣٤	٤٠	٤٦	٥٢
عدد ساعات التمرينات (ص)	٨	٧	٥	٢	٣	١,٥	١

أ ارسم مخطط الانتشار.

ب حدّد نوع العلاقة.

الحل:

أ



ب من مخطط الانتشار نلاحظ أنه إذا زادت قيمة س

تتناقص تبعًا لها قيمة ص

∴ الارتباط عكسي (سالب)

العلاقة خطية

حاول أن تحل

٣ ارسم مخطط الانتشار للبيانات التالية وحدّد نوع العلاقة التي تعبر عنها:

س	٢	٣	٤	٥	٦	٧
ص	٧	٥	٤	٣	٢	١

(٢–١–ب) مُعامل الارتباط الخطي Linear Correlation Coefficient

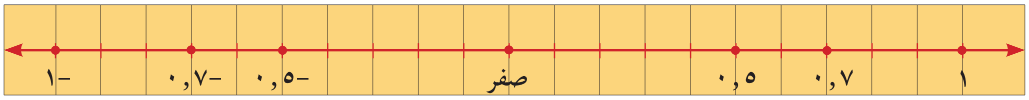
نعلم أن الاستنتاجات المبنيّة على المعاينات البصريّة لمخطط الانتشار هي نسبيّة بامتياز، لذا فنحن بحاجة إلى قياسات أكثر دقة وموضوعية بالتالي نستخدم مُعامل الارتباط الخطّي (ر).

تعريف: مُعامل الارتباط الخطّي (ر)

هو عبارة عن مقياس عددي لقوّة العلاقة بين متغيّرين يمثّلان بيانات كمّية،

حيث −١ ≤ ر ≤ ١.

خواص مُعامل الارتباط (ر)



١ -١ ≤ ر ≤ ١ أو ر ∈ [-١ ، ١].

٢ إذا كانت ر = ١ يكون الارتباط طردي (موجب) تام.

٣ إذا كانت ر = -١ يكون الارتباط عكسي (سالب) تام.

٤ إذا كانت ر = ٠ ينعدم الارتباط.

٥ إذا كانت ر ∈ [٠,٧ ، ١) يكون الارتباط طردي (موجب) قوي.

٦ إذا كانت ر ∈ [٠,٥ ، ٠,٧) يكون الارتباط طردي (موجب) متوسط.

٧ إذا كانت ر ∈ (٠ ، ٠,٥) يكون الارتباط طردي (موجب) ضعيف.

٨ إذا كانت ر ∈ (-٠,٥ ، ٠) يكون الارتباط عكسي (سالب) ضعيف.

٩ إذا كانت ر ∈ (-٠,٧ ، -٠,٥] يكون الارتباط عكسي (سالب) متوسط.

١٠ إذا كانت ر ∈ (-١ ، -٠,٧] يكون الارتباط عكسي (سالب) قوي.

مُعامل ارتباط بيرسون Pearson Correlation Coefficient

$$ر = \frac{\sum(س - \overline{س})(ص - \overline{ص})}{ن\, ع_{س}\, ع_{ص}}$$ حيث:

$$ع_{س} = \sqrt{\frac{\sum(س - \overline{س})^{٢}}{ن}}$$ حيث $ع_{س}$ (الانحراف المعياري للمتغير س)

$$ع_{ص} = \sqrt{\frac{\sum(ص - \overline{ص})^{٢}}{ن}}$$ حيث $ع_{ص}$ (الانحراف المعياري للمتغير ص)

$$ر = \frac{\sum(س - \overline{س})(ص - \overline{ص})}{\sqrt{\sum(س - \overline{س})^{٢}}\sqrt{\sum(ص - \overline{ص})^{٢}}}$$

مثال (٤)

س	٧	٤	٢	١	١
ص	٢٣	١٥	٨	٥	٤

من الجدول المقابل:

أ أوجد مُعامل الارتباط ر.

ب حدّد نوع وقوة الارتباط.

الحل:

أ مُعامل الارتباط:

$$\text{ر} = \frac{\sum(\text{س} - \bar{\text{س}})(\text{ص} - \bar{\text{ص}})}{\sqrt{\sum(\text{س} - \bar{\text{س}})^{٢}} \sqrt{\sum(\text{ص} - \bar{\text{ص}})^{٢}}}$$

س	ص	س $-\bar{\text{س}}$	ص $-\bar{\text{ص}}$	(س $-\bar{\text{س}}$)^{٢}	(ص $-\bar{\text{ص}}$)^{٢}	(س $-\bar{\text{س}}$)(ص $-\bar{\text{ص}}$)
٧	٢٣	٤	١٢	١٦	١٤٤	٤٨
٤	١٥	١	٤	١	١٦	٤
٢	٨	١-	٣-	١	٩	٣
١	٥	٢-	٦-	٤	٣٦	١٢
١	٤	٢-	٧-	٤	٤٩	١٤
المجموع $\sum$س = ١٥	$\sum$ص = ٥٥			$\sum$(س $-\bar{\text{س}}$)^{٢} = ٢٦	$\sum$(ص $-\bar{\text{ص}}$)^{٢} = ٢٥٤	$\sum$(س $-\bar{\text{س}}$)(ص $-\bar{\text{ص}}$) = ٨١

$\because$ $\bar{\text{س}} = \frac{\sum \text{س}}{\text{ن}} = \frac{١٥}{٥} = ٣$ ، $\bar{\text{ص}} = \frac{\sum \text{ص}}{\text{ن}} = \frac{٥٥}{٥} = ١١$

$\therefore$ مُعامل الارتباط $= \frac{٨١}{\sqrt{٢٦} \times \sqrt{٢٥٤}} \approx ٠,٩٩٦٨$

ب نوع الارتباط: طردي موجب قوي.

حاول أن تحل

٤ يبيّن الجدول التالي العلاقة بين وزن مولود جديد وطوله خلال فترة محددة من الزمن.

الوزن (كجم)	٢,١	٢,٩	٣,٢	٣,٨	٤,١
الطول (سم)	٥٨	٦٥	٦٨	٧١	٧٥

أ أوجد مُعامل الارتباط ر.

ب حدّد نوع وقوة الارتباط.

مثال (٥)

أوجد مُعامل الارتباط ر وحدّد نوعه وقوته للمتغيرين س ، ص حيث :

س	١	٢	٣	٤	٥
ص	١	−١	−٤	−٦	−٥

الحل:

مُعامل الارتباط: $\text{ر} = \dfrac{\sum(\text{س} - \overline{\text{س}})(\text{ص} - \overline{\text{ص}})}{\sqrt{\sum(\text{س} - \overline{\text{س}})^{2}}\ \sqrt{\sum(\text{ص} - \overline{\text{ص}})^{2}}}$

س	ص	$\text{س} - \overline{\text{س}}$	$\text{ص} - \overline{\text{ص}}$	$(\text{س} - \overline{\text{س}})^{2}$	$(\text{ص} - \overline{\text{ص}})^{2}$	$(\text{س} - \overline{\text{س}})(\text{ص} - \overline{\text{ص}})$
١	١	−٢	٤	٤	١٦	−٨
٢	−١	−١	٢	١	٤	−٢
٣	−٤	٠	−١	٠	١	٠
٤	−٦	١	−٣	١	٩	−٣
٥	−٥	٢	−٢	٤	٤	−٤
المجموع $\sum\text{س} = ١٥$	$\sum\text{ص} = −١٥$			$\sum(\text{س} - \overline{\text{س}})^{2} = ١٠$	$\sum(\text{ص} - \overline{\text{ص}})^{2} = ٣٤$	$\sum(\text{س} - \overline{\text{س}})(\text{ص} - \overline{\text{ص}}) = −١٧$

$\because\ \overline{\text{س}} = \dfrac{\sum\text{س}}{\text{ن}} = \dfrac{١٥}{٥} = ٣$ ، $\overline{\text{ص}} = \dfrac{\sum\text{ص}}{\text{ن}} = \dfrac{−١٥}{٥} = −٣$

$\therefore$ مُعامل الارتباط: $\text{ر} = \dfrac{−١٧}{\sqrt{١٠} \times \sqrt{٣٤}} \approx −0,9220$

نوع الارتباط: عكسي سالب قوي.

حاول أن تحل

٥ أوجد مُعامل الارتباط ر وحدد نوعه وقوته للمتغيرين س ، ص حيث:

س	٨	١٠	٦	٤	١٥	١٣	٥	١١	٩
ص	١٥٠	١٦٠	١٥٠	١٣٠	١٦٠	١٨٠	١٢٠	١٦٠	١٥٠

مثال (٦)

احسب مُعامل الارتباط الخطي للبيانات التالية وحدّد نوعه وقوته.

س	١	٢	٣	٤	٥
ص	٣	٥	٧	٩	١١

الحل:

مُعامل الارتباط: $ر = \dfrac{\Sigma(\text{س} - \overline{\text{س}})(\text{ص} - \overline{\text{ص}})}{\sqrt{\Sigma(\text{س} - \overline{\text{س}})^{٢}}\ \sqrt{\Sigma(\text{ص} - \overline{\text{ص}})^{٢}}}$

	س	ص	$\text{س} - \overline{\text{س}}$	$\text{ص} - \overline{\text{ص}}$	$(\text{س} - \overline{\text{س}})^{٢}$	$(\text{ص} - \overline{\text{ص}})^{٢}$	$(\text{س} - \overline{\text{س}})(\text{ص} - \overline{\text{ص}})$
	١	٣	−٢	−٤	٤	١٦	٨
	٢	٥	−١	−٢	١	٤	٢
	٣	٧	٠	٠	٠	٠	٠
	٤	٩	١	٢	١	٤	٢
	٥	١١	٢	٤	٤	١٦	٨
المجموع	$\Sigma\text{س} = ١٥$	$\Sigma\text{ص} = ٣٥$			$\Sigma(\text{س} - \overline{\text{س}})^{٢} = ١٠$	$\Sigma(\text{ص} - \overline{\text{ص}})^{٢} = ٤٠$	$\Sigma(\text{س} - \overline{\text{س}})(\text{ص} - \overline{\text{ص}}) = ٢٠$

$$\overline{\text{س}} = \frac{\Sigma\text{س}}{\text{ن}} = ٣ \ ،\ \overline{\text{ص}} = \frac{\Sigma\text{ص}}{\text{ن}} = ٧$$

$$\therefore\ ر = \frac{٢٠}{\sqrt{١٠} \times \sqrt{٤٠}}$$

$$= ١$$

نوع الارتباط: طردي (موجب) تام.

حاول أن تحل

٦ احسب مُعامل الارتباط الخطي للبيانات التالية وحدّد نوعه وقوته.

س	١	٢	٣	٤	٥
ص	٤	٣	٢	١	٠

صيغة أخرى لمعامل ارتباط بيرسون

$$ر = \frac{ن(\sum س\,ص) - (\sum س)(\sum ص)}{\sqrt{ن(\sum س^{2}) - (\sum س)^{2}}\ \sqrt{ن(\sum ص^{2}) - (\sum ص)^{2}}}$$

مثال (٧)

يبيّن الجدول التالي العلاقة بين أطوال عدد من الدببة وأوزانها، وذلك ضمن فترة محددة من أعمارها.

الطول (سم)	١٣٥	١٧٠	١٨٠	١٨٢	١٨٧	١٧٤	١٨٥	٩٤
الوزن (كجم)	٣٦	١٥٦	١٨٨	١٥٨	١١٩	١٦٣	١٥٠	١٥

استخدم الجدول أعلاه لإيجاد مُعامل الارتباط الخطي ر والذي يحدّد العلاقة بين أطوال الدببة وأوزانها ثم بيّن نوعه وقوته.

الحل:

$$ر = \frac{ن(\sum س\,ص) - (\sum س)(\sum ص)}{\sqrt{ن(\sum س^{2}) - (\sum س)^{2}}\ \sqrt{ن(\sum ص^{2}) - (\sum ص)^{2}}}$$

	س (الطول)	ص (الوزن)	س ص	$س^{2}$	$ص^{2}$
	١٣٥	٣٦	٤٨٦٠	١٨٢٢٥	١٢٩٦
	١٧٠	١٥٦	٢٦٥٢٠	٢٨٩٠٠	٢٤٣٣٦
	١٨٠	١٨٨	٣٣٨٤٠	٣٢٤٠٠	٣٥٣٤٤
	١٨٢	١٥٨	٢٨٧٥٦	٣٣١٢٤	٢٤٩٦٤
	١٨٧	١١٩	٢٢٢٥٣	٣٤٩٦٩	١٤١٦١
	١٧٤	١٦٣	٢٨٣٦٢	٣٠٢٧٦	٢٦٥٦٩
	١٨٥	١٥٠	٢٧٧٥٠	٣٤٢٢٥	٢٢٥٠٠
	٩٤	١٥	١٤١٠	٨٨٣٦	٢٢٥
المجموع	$\sum$ س = ١٣٠٧	$\sum$ ص = ٩٨٥	$\sum$ س ص = ١٧٣٧٥١	$\sum س^{2}$ = ٢٢٠٩٥٥	$\sum ص^{2}$ = ١٤٩٣٩٥

$$ر = \frac{٨ (١٧٣٧٥١) - (١٣٠٧) (٩٨٥)}{\sqrt{٨ (٢٢٠٩٥٥) - (١٣٠٧)^٢} \sqrt{٨ (١٤٩٣٩٥) - (٩٨٥)^٢}}$$

$$ر = \frac{١٠٢ ٦١٣}{١١٥ ٥٨٢}$$

$$ر \approx ٠,٨٨٧٨$$

نوع الارتباط: طردي (موجب) قوي.

حاول أن تحل

٧ من الجدول التالي احسب مُعامل الارتباط الخطي وبيّن نوعه وقوته.

س	١	٢	٣	٤	٥	٦
ص	٥٩	٦٥	٧٠	٧٢	٨٠	٥٢

مثال (٨)

احسب مُعامل الارتباط الخطي للمتغيرين التاليين وبيّن نوعه وقوته.

س	١	٢	٣	٤	٥	٦
ص	٤	٧	٨	٣	٥	٥

الحل:

$$ر = \frac{ن (\sum س ص) - (\sum س) (\sum ص)}{\sqrt{ن (\sum س^٢) - (\sum س)^٢} \sqrt{ن (\sum ص^٢) - (\sum ص)^٢}}$$

	س	ص	س ص	$س^٢$	$ص^٢$
	١	٤	٤	١	١٦
	٢	٧	١٤	٤	٤٩
	٣	٨	٢٤	٩	٦٤
	٤	٣	١٢	١٦	٩
	٥	٥	٢٥	٢٥	٢٥
	٦	٥	٣٠	٣٦	٢٥
المجموع	$\sum س = ٢١$	$\sum ص = ٣٢$	$\sum س ص = ١٠٩$	$\sum س^٢ = ٩١$	$\sum ص^٢ = ١٨٨$

$$ر = \frac{٦ \times ١٠٩ - ٢١ \times ٣٢}{\sqrt{٦ \times ٩١ - (٢١)^{٢}} \; \sqrt{٦ \times ١٨٨ - (٣٢)^{٢}}}$$

$$ر = \frac{-١٨}{\sqrt{١٠٥} \times \sqrt{١٠٤}}$$

$$ر \approx -٠,١٧٢٣$$

نوع الارتباط: عكسي (سالب) ضعيف.

حاول أن تحل

٨ احسب مُعامل الارتباط الخطي للبيانات التالية وبيّن نوعه وقوته:

س	٢	٣	٤	٤	٥	٦
ص	٩٨	٩٩	٧٥	٤٠	١٠٠	١٥٠

مثال (٩)

احسب مُعامل الارتباط الخطي للبيانات التالية وبيّن نوعه وقوته.

س	٨	١٥	١٠	١٤	٩	١٢	١٣	١١
ص	٨	١	٦	٢	٧	٤	٣	٥

الحل:

$$ر = \frac{ن(\sum س\,ص) - (\sum س)(\sum ص)}{\sqrt{ن(\sum س^{٢}) - (\sum س)^{٢}} \; \sqrt{ن(\sum ص^{٢}) - (\sum ص)^{٢}}}$$

	س	ص	س ص	$س^{٢}$	$ص^{٢}$
	٨	٨	٦٤	٦٤	٦٤
	١٥	١	١٥	٢٢٥	١
	١٠	٦	٦٠	١٠٠	٣٦
	١٤	٢	٢٨	١٩٦	٤
	٩	٧	٦٣	٨١	٤٩
	١٢	٤	٤٨	١٤٤	١٦
	١٣	٣	٣٩	١٦٩	٩
	١١	٥	٥٥	١٢١	٢٥
المجموع	$\sum س = ٩٢$	$\sum ص = ٣٦$	$\sum س\,ص = ٣٧٢$	$\sum س^{٢} = ١١٠٠$	$\sum ص^{٢} = ٢٠٤$

$$ر = \frac{٨ \times ٣٧٢ - ٩٢ \times ٣٦}{\sqrt{٨ \times ١١٠٠ - (٩٢)^{٢}} \times \sqrt{٨ \times ٢٠٤ - (٣٦)^{٢}}}$$

$$ر = \frac{-٣٣٦}{\sqrt{٣٣٦} \times \sqrt{٣٣٦}}$$

$$ر = \frac{-٣٣٦}{٣٣٦}$$

$$ر = -١$$

نوع الارتباط: عكسي (سالب) تام

حاول أن تحل

٩ احسب مُعامل الارتباط الخطي للبيانات التالية وبيّن نوعه وقوته.

س	٨	٥	١١	٧	٩	١٢	٦
ص	٤	١	٧	٣	٥	٨	٢

مثال (١٠)

في ما يلي درجات عدد من الطلاب في مادتي الإحصاء (س) والتاريخ (ص)

الإحصاء (س)	٥	١٠	٦	١٥	١١	١٣	١٧	١٢
التاريخ (ص)	١٧	١٧	١٥	٩	١٠	١٠	٦	١٢

أ أوجد مُعامل الارتباط ر.

ب حدّد نوع وقوة الارتباط.

الحل:

أ $ر = \frac{ن(\sum س ص) - (\sum س)(\sum ص)}{\sqrt{ن(\sum س^٢) - (\sum س)^٢} \times \sqrt{ن(\sum ص^٢) - (\sum ص)^٢}}$

س	ص	س ص	س²	ص²
٥	١٧	٨٥	٢٥	٢٨٩
١٠	١٧	١٧٠	١٠٠	٢٨٩
٦	١٥	٩٠	٣٦	٢٢٥
١٥	٩	١٣٥	٢٢٥	٨١
١١	١٠	١١٠	١٢١	١٠٠
١٣	١٠	١٣٠	١٦٩	١٠٠
١٧	٦	١٠٢	٢٨٩	٣٦
١٢	١٢	١٤٤	١٤٤	١٤٤
المجموع $\sum س = ٨٩$	$\sum ص = ٩٦$	$\sum س ص = ٩٦٦$	$\sum س^٢ = ١١٠٩$	$\sum ص^٢ = ١٢٦٤$

$ر = \frac{٨(٩٦٦) - (٨٩)(٩٦)}{\sqrt{٨(١١٠٩) - (٨٩)^٢} \times \sqrt{٨(١٢٦٤) - (٩٦)^٢}}$

$ر = \frac{٧٧٢٨ - ٨٥٤٤}{\sqrt{٩٥١} \times \sqrt{٨٩٦}} = \frac{-٨١٦}{\sqrt{٩٥١} \times \sqrt{٨٩٦}}$

$ر \approx -٠{,}٨٨٤٠$

ب نوع الارتباط: عكسي (سالب) قوي.

حاول أن تحل

١٠ أوجد مُعامل الارتباط الخطي للبيانات التالية وحدّد نوعه وقوته.

س	١	٣	٨	٧	٦	٥	٧	٨
ص	١٩	١٦	١٦	١٩	١٨	١٧	١١	٩

الانحدار

Regression

سوف تتعلم

- إيجاد معادلة خط الانحدار.
- التقدير باستخدام معادلة خط الانحدار.
- إيجاد مقدار الخطأ.

دعنا نفكر ونتناقش

في الجدول التالي قيم لمتغيّرين: طول الأم (س) وطول ابنتها (ص) بالسنتيمتر.

طول الأم (س)	١٦٠	١٦٨	١٦٩	١٦٤	١٧٤	١٦١	١٦٦	١٥٨
طول الابنة (ص)	١٥٨	١٦٧	١٧٠	١٦٣	١٧١	١٦٥	١٧٢	١٥٧

لدينا ر ≈ ٠,٨٤٤، إذًا يوجد علاقة خطيّة طرديّة قوية بين طول الأم وطول ابنتها.

أضفنا زوج المتغيرين ($\overline{س}$ ، $\overline{ص}$) = (١٦٥، ١٦٥,٣٧٥) إلى الجدول حيث $\overline{س}$ = ١٦٥هو المتوسط الحسابي لأطوال الأمهات، $\overline{ص}$ = ١٦٥,٣٧٥ هو المتوسط الحسابي لأطوال البنات فلاحظنا أن قيمة ر لم تتغير.

نريد أن نقدّر طول الابنة من خلال العلاقة مع طول أمها، لذا افترضنا زوج المتغيرين (١٥٠، ١٧٠) وأضفناه إلى الجدول.

- هل يتوافق زوج المتغيرين الذي أضفناه مع الجدول علمًا أنّ قيمة ر تصبح ٠,٢١٦؟
- هل يمكن التنبؤ بقيمة إحدى الظاهرتين إذا علمت قيمة الظاهرة الأخرى؟ وكيف؟

الانحدار Regression

لقد تعلمنا في الدرس السابق مفهوم الارتباط والارتباط الخطي، وعرفنا كيف يمكن حساب قيمة مُعامل الارتباط الخطي بين متغيرين، وعليه تمّ تحديد قوة العلاقة الخطية بين المتغيرين ونوع هذه العلاقة فيما إذا كانت طردية أم عكسية.

وفي هذا الدرس سوف نتعلّم وصف العلاقة بين متغيّرين بإيجاد معادلة الخط المستقيم الممثّل لهذه العلاقة.

يسمّى هذا الخط المستقيم بخط الانحدار، وتسمّى معادلته بمعادلة خط الانحدار.

تعريف: الانحدار

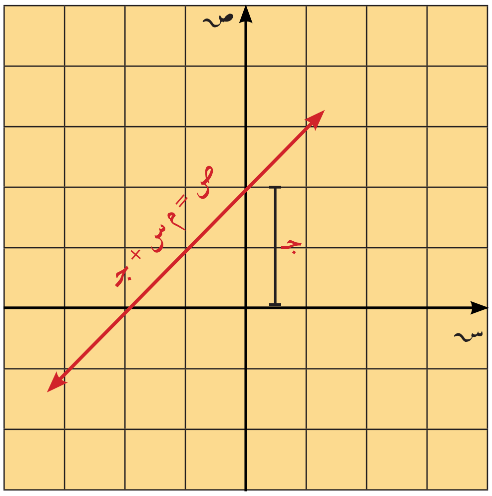
هو وصف العلاقة بين متغيرين.

معادلة خط الانحدار — Equation of Linear Regression

تعريف: معادلة خط الانحدار

هي المعادلة الخطية التي يمكن من خلالها التنبؤ بقيمة أحد المتغيرين إذا علمت قيمة المتغير الآخر.

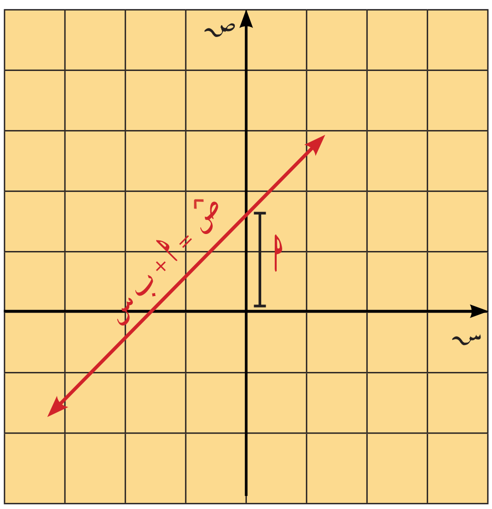
سبق لنا دراسة معادلة الخط المستقيم على الصورة: ص = م س + جـ ، حيث م ترمز إلى ميل هذا المستقيم، |جـ| ترمز إلى طول الجزء المقطوع من محور الصادات.



شكل (١)

أما في الإحصاء معادلة خط انحدار مستقيم تكتب على الصورة:

صَ = أ + ب س، حيث |أ| ترمز إلى طول الجزء المقطوع من محور الصادات، ب ترمز إلى ميل المستقيم.



شكل (٢)

حيث: $ب = \frac{ن\,(\sum س\,ص) - (\sum س)(\sum ص)}{ن\,(\sum س^{٢}) - (\sum س)^{٢}}$

$أ = \overline{ص} - ب\,\overline{س}$

حيث: $\overline{ص} = \frac{\sum ص}{ن}$ ، $\overline{س} = \frac{\sum س}{ن}$

خطوات إيجاد معادلة خط الانحدار: $\hat{ص}$ = أ + ب س

١ تعيين قيمة ب

٢ تعيين قيمة أ

٣ نكتب معادلة خط الانحدار: $\hat{ص}$ = أ + ب س

٤ التنبؤ بقيمة ص إذا علمت قيمة س

٥ تحديد مقدار الخطأ في التنبؤ.

مقدار الخطأ = |القيمة الجدولية − القيمة التي تحقق معادلة الانحدار|

$= |ص_{س} - \hat{ص}_{س}|$

مثال (١)

سقطت كرة من ارتفاع ٥٠ مترًا، وتمّ تسجيل المسافات (بالمتر) التي قطعتها هذه الكرة كل ٠,٥ ثانية لمدة ثلاث ثوان.

فأتت النتائج كما يوضح الجدول التالي:

الوقت (س)	٠	٠,٥	١	١,٥	٢	٢,٥	٣
المسافة (ص)	٠	١,٢	٤,٩	١١	١٩,٥	٣٠,٥	٤٤

أ أوجد معادلة خط الانحدار.

ب قدّر قيمة المسافة ص عندما س = ٤

جـ أوجد مقدار الخطأ في المسافة عندما س = ٢,٥ ثانية.

الحل:

$$ب = \frac{ن (\sum س ص) - (\sum س)(\sum ص)}{ن (\sum س^{٢}) - (\sum س)^{٢}}$$

أ

	س	ص	س ص	$س^{٢}$
	٠	٠	٠	٠
	٠,٥	١,٢	٠,٦	٠,٢٥
	١	٤,٩	٤,٩	١
	١,٥	١١	١٦,٥	٢,٢٥
	٢	١٩,٥	٣٩	٤
	٢,٥	٣٠,٥	٧٦,٢٥	٦,٢٥
	٣	٤٤	١٣٢	٩
المجموع	$\sum$س = ١٠,٥	$\sum$ص = ١١١,١	$\sum$س ص = ٢٦٩,٢٥	$\sum س^{٢}$ = ٢٢,٧٥

ن = ٧ ، $\bar{س} = \frac{١٠,٥}{٧}$ = ١,٥ ، $\bar{ص} = \frac{١١١,١}{٧}$ = ١٥,٨٧١٤

$$ب = \frac{٧ \times ٢٦٩,٢٥ - ١٠,٥ \times ١١١,١}{٧ \times ٢٢,٧٥ - (١٠,٥)^{٢}}$$

$$ب = \frac{٧١٨,٢}{٤٩}$$

ب ≈ ١٤,٦٥٧١

أ = $\bar{ص}$ − ب $\bar{س}$

= ١٥,٨٧١٤ − ١٤,٦٥٧١ × ١,٥

≈ −٦,١١٤٣

معادلة خط الانحدار هي: $\hat{ص}$ = −٦,١١٤٣ + ١٤,٦٥٧١س

ب ∵ $\hat{ص}$ = −٦,١١٤٣ + ١٤,٦٥٧١ س

∴ المسافة ص عندما س = ٤ هي:

$\hat{ص}_{٤}$ = −٦,١١٤٣ + ١٤,٦٥٧١ × ٤

≈ ٥٢,٥١٤١ مترًا.

جـ عند س = ٢,٥ ثانية   من المعادلة $\hat{ص}$ = -٦,١١٤٣ + ١٤,٦٥٧١ س

نجد أن $\hat{ص}_{٢,٥}$ = -٦,١١٤٣ + ١٤,٦٥٧١ × ٢,٥

≈ ٣٠,٥٢٨٥

من الجدول عند س = ٢,٥ ثانية نجد أن ص = ٣٠,٥

∴ مقدار الخطأ = |$ص_{س}$ - $\hat{ص}_{س}$|

= |٣٠,٥ - ٣٠,٥٢٨٥|

= ٠,٠٢٨٥

حاول أن تحل

١ في الجدول التالي: المتغير س هو تكلفة إنتاج فيلم سينمائي (بملايين الدولارات) والمتغيّر ص هو مردود هذا الفيلم.

التكلفة (س)	٦٢	٩٠	٥٠	٣٥	٢٠٠	١٠٠	٩٥
المردود (ص)	٦٥	٦٤	٤٨	٥٧	٦٠١	١٤٦	٤٧

أ أوجد معادلة خط الانحدار.

ب قدّر مردود فيلم بلغت تكلفته ٥٥ مليون دولار.

جـ أوجد مقدار الخطأ لفيلم بلغت تكلفته ٩٠ مليون دولار.

مثال (٢)

من الجدول التالي:

س	٤٣	٤٨	٥٦	٦١	٦٧	٧٠
ص	١٢٨	١٢٠	١٣٥	١٤٣	١٤١	١٥٢

أوجد:

أ معادلة خط الانحدار.

ب قيمة ص عندما س = ٥٢

جـ مقدار الخطأ عندما س = ٦٧

الحل: $\text{ب} = \dfrac{\text{ن}\,(\sum \text{س ص}) - (\sum \text{س})(\sum \text{ص})}{\text{ن}\,(\sum \text{س}^{\text{٢}}) - (\sum \text{س})^{\text{٢}}}$

أ

	س	ص	س ص	س٢
	٤٣	١٢٨	٥ ٥٠٤	١ ٨٤٩
	٤٨	١٢٠	٥ ٧٦٠	٢ ٣٠٤
	٥٦	١٣٥	٧ ٥٦٠	٣ ١٣٦
	٦١	١٤٣	٨ ٧٢٣	٣ ٧٢١
	٦٧	١٤١	٩ ٤٤٧	٤ ٤٨٩
	٧٠	١٥٢	١٠ ٦٤٠	٤ ٩٠٠
المجموع	$\sum$ س = ٣٤٥	$\sum$ ص = ٨١٩	$\sum$ س ص = ٤٧ ٦٣٤	$\sum$ س٢ = ٢٠ ٣٩٩

ن = ٦ ، $\overline{\text{س}} = \dfrac{\text{٣٤٥}}{\text{٦}}$ = ٥٧,٥ ، $\overline{\text{ص}} = \dfrac{\text{٨١٩}}{\text{٦}}$ = ١٣٦,٥

$\text{ب} = \dfrac{\text{٦} \times \text{٤٧ ٦٣٤} - \text{٣٤٥} \times \text{٨١٩}}{\text{٦} \times \text{٢٠ ٣٩٩} - (\text{٣٤٥})^{\text{٢}}}$

≈ ٠,٩٦٤٤

أ = $\overline{\text{ص}}$ − ب $\overline{\text{س}}$

= ١٣٦,٥ − ٠,٩٦٤٤ × ٥٧,٥

= ٨١,٠٤٧٠

∴ معادلة خط الانحدار هي: $\hat{\text{ص}}$ = ٨١,٠٤٧٠ + ٠,٩٦٤٤ س

ب عندما س = ٥٢

$\hat{\text{ص}}$ = ٨١,٠٤٧٠ + ٠,٩٦٤٤ × ٥٢

≈ ١٣١,١٩٥٨

جـ من الجدول عند س = ٦٧ نوجد ص = ١٤١

$\hat{\text{ص}}_{\text{٦٧}}$ = ٨١,٠٤٧٠ + ٠,٩٦٤٤ × ٦٧

≈ ١٤٥,٦٦١٨

∴ مقدار الخطأ = $|ص_س - \hat{ص}_س|$

$= |١٤١ - ١٤٥,٦٦١٨| = ٤,٦٦١٨$

حاول أن تحل

٢ من الجدول التالي:

س	١٨٠	٢٠٥	١٩٧	١٩٢	١٨٩	٢٠٣	١٩٧	١٨٤
ص	٨٥	١١٧	٨٢	٩٧	٩٢	٨٠	١١٠	١٢٢

أوجد:

أ معادلة خط الانحدار.

ب قيمة ص عندما س = ٢٠٠

جـ مقدار الخطأ عندما س = ١٩٢

مثال (٣)

باستخدام البيانات التالية لقيم س ، ص.

س	١	٣	٥	٧	٩
ص	٢	٥	٩	١٠	١٤

أوجد:

أ معادلة خط الانحدار.

ب قيمة ص عندما س = ١٠

جـ مقدار الخطأ عندما س = ٥

الحل:

$$ب = \frac{ن (\sum س ص) - (\sum س)(\sum ص)}{ن (\sum س^٢) - (\sum س)^٢}$$

أ

	س	ص	س ص	$س^٢$
	١	٢	٢	١
	٣	٥	١٥	٩
	٥	٩	٤٥	٢٥
	٧	١٠	٧٠	٤٩
	٩	١٤	١٢٦	٨١
المجموع	$\sum س = ٢٥$	$\sum ص = ٤٠$	$\sum س ص = ٢٥٨$	$\sum س^٢ = ١٦٥$

ن = ٥ ، $\overline{\text{س}} = \frac{\text{مجـ س}}{\text{ن}} = \frac{٢٥}{٥} = ٥$ ، $\overline{\text{ص}} = \frac{\text{مجـ ص}}{\text{ن}} = \frac{٤٠}{٥} = ٨$

$\text{ب} = \frac{٥ \times ٢٥٨ - ٢٥ \times ٤٠}{٥ \times ١٦٥ - ٢٥ \times ٢٥} = ١{,}٤٥$

$\text{أ} = \overline{\text{ص}} - \text{ب}\ \overline{\text{س}}$

$= ٨ - ١{,}٤٥ \times ٥$

$= ٠{,}٧٥$

∴ معادلة خط الانحدار هي: $\hat{\text{ص}} = \text{أ} + \text{ب س}$

$\hat{\text{ص}} = ٠{,}٧٥ + ١{,}٤٥\ \text{س}$

ب عندما س = ١٠ فإن:

ص = ٠,٧٥ + ١,٤٥ × ١٠ = ١٥,٢٥

جـ من الجدول ص = ٩

من المعادلة $\hat{\text{ص}} = ٠{,}٧٥ + ١{,}٤٥ \times ٥ = ٨$

∴ مقدار الخطأ = $|\text{ص}_{٥} - \hat{\text{ص}}_{٥}|$

$= |٩ - ٨| = ١$

حاول أن تحل

٣ من الجدول التالي:

س	٤	٥	٨	٩	١٠	١٢
ص	٢	٤	٥	٨	٦	١١

أوجد:

أ معادلة خط الانحدار.

ب قيمة ص عندما س = ١٠

جـ أوجد مقدار الخطأ عندما س = ١٠

المرشد لحل المسائل

في متجر للأدوات الكهربائية ، تختلف أسعار آلات التصوير الرقميّة بحسب نقاوة صورتها التي تقاس بالميغابيكسل.

يوضّح الجدول التالي أسعار إحدى هذه الآلات ومدى نقاوة صورتها:

(س) نقاوة الصورة بالميغابيكسل	٥	٨	١٠	١٤	١٦	١٨
(ص) السعر بالدينار الكويتي	٢٥	٣٥	٥٠	٨٥	١٢٠	١٤٠

أراد جاسم تقدير سعر آلة نقاوتها ٢٠ ميغابيكسل، علمًا أنّه سمع من صاحب المتجر أنّه يوجد علاقة بين السعر والنقاوة.

فكّر جاسم:

إذا رسمت مخطط الانتشار للأسعار والنقاوة ، أتعرّف على طبيعة هذه العلاقة.

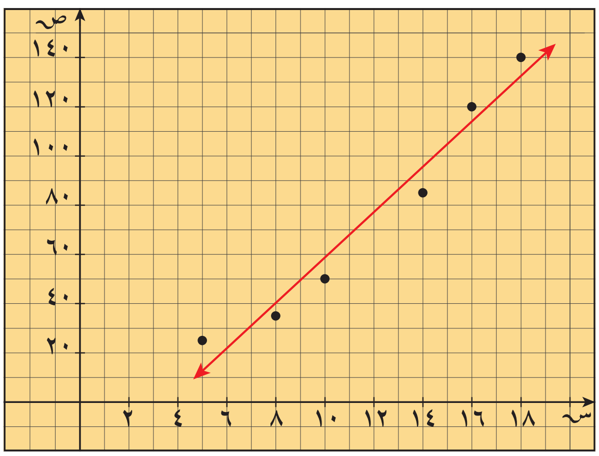
فلاحظ أن هذه العلاقة هي خطيّة طرديّة ، لذا أراد إيجاد قيمة مُعامل الارتباط الخطي ومعادلة خط الانحدار.

أوجد القيم التالية التي ستساعده على ذلك:

ن = ٦ ، $\sum$ س = ٧١ ، $\sum$ ص = ٤٥٥ ، $\sum$ س ص = ٦٥٣٥ ،

$\sum$ س^{٢} = ٩٦٥ ، $\sum$ ص^{٢} = ٤٥ ٥٧٥ ، ($\sum$ س)^{٢} = ٥٠٤١ ، ($\sum$ ص)^{٢} = ٢٠٧ ٠٢٥ ،

$\overline{\text{س}}$ = ١١,٨ ، $\overline{\text{ص}}$ = ٧٥,٨



قيمة مُعامل الارتباط الخطي ر ≈ ٠,٩٧٨٨ ، وهذا يدلّ على علاقة خطيّة قويّة بين السعر والنقاوة.

معادلة خط الانحدار: $\hat{\text{ص}}$ = −٣٣ + ٩,٢٢ س

لتقدير سعر آلة مع ٢٠ ميغابيكسل، نعوّض س = ٢٠

ونحصل على $\hat{\text{ص}}$ ≈ ١٥١ دينارًا كويتيًّا.

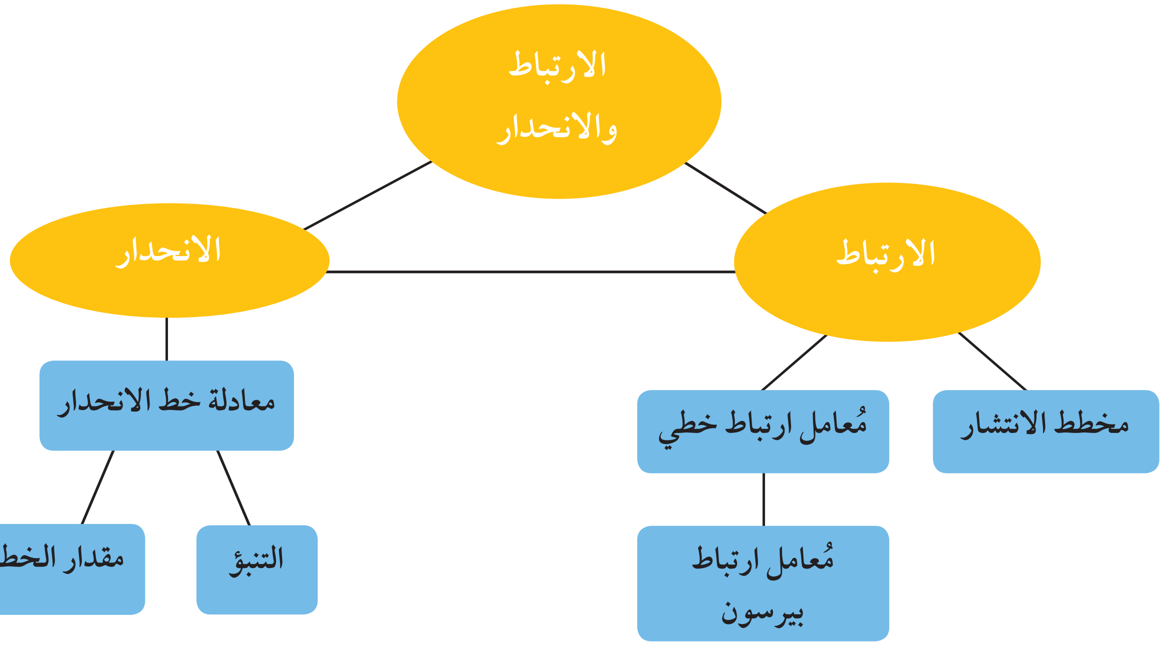
مسألة إضافية

أجري في المتجر نفسه تخفيض على الأسعار بنسبة ١٥٪ .

برأيك، كيف يتغير تقدير جاسم؟ أعد الحل مستخدمًا السعر المخفّض للتأكّد من إجابتك.

(**ملاحظة**: استخدم الجدول نفسه من المسألة السابقة إنّما بتخفيض قدره ١٥٪ على الأسعار)

مخطط تنظيمي للوحدة الثانية



ملخص

- الارتباط هو طريقة إحصائية يمكن من خلالها تحديد العلاقة بين متغيرين.
- مخطط الانتشار هو شكل بياني لعدد من الأزواج المرتبة (س، ص) يستخدم لوصف العلاقة الموجودة بين متغيرين.
- العلاقة بين متغيّرين تكون:
  - علاقة خطيّة طردية: تنتشر النقاط على جانبي خط مستقيم تصاعديّ.
  - علاقة خطيّة عكسيّة: تنتشر النقاط على جانبي خط مستقيم تنازليّ.
  - علاقة غير خطيّة: تنتشر النقاط على جانبي خط منحن.
  - لا توجد علاقة: لا يوجد نمط محدد لانتشار النقاط في الشكل البياني.
- مُعامل الارتباط الخطي يقيس قوة العلاقة الخطيّة بين متغيرين متصلين ونوعها،

$$\text{ر} = \frac{\text{ن}\,\text{مج}\,\text{س ص} - \text{مج}\,\text{س}\,\text{مج}\,\text{ص}}{\sqrt{\text{ن}\,\text{مج}\,\text{س}^2 - (\text{مج}\,\text{س})^2}\,\sqrt{\text{ن}\,\text{مج}\,\text{ص}^2 - (\text{مج}\,\text{ص})^2}} \quad \text{أو} \quad \text{ر} = \frac{\text{مج}(\text{س} - \overline{\text{س}})\,\text{مج}(\text{ص} - \overline{\text{ص}})}{\sqrt{\text{مج}(\text{س} - \overline{\text{س}})^2}\,\sqrt{\text{مج}(\text{ص} - \overline{\text{ص}})^2}}$$

- الانحدار هو طريقة إحصائية تستخدم لوصف طبيعة العلاقة بين متغيرين س ، ص من حيث كونها خطيّة أو غير خطيّة.
- معادلة خط الانحدار $\hat{\text{ص}} = \text{أ} + \text{ب س}$ ، حيث:

$$\text{ب} = \frac{\text{ن}(\text{مج}\,\text{س ص}) - (\text{مج}\,\text{س})(\text{مج}\,\text{ص})}{\text{ن}(\text{مج}\,\text{س}^2) - (\text{مج}\,\text{س})^2}$$

$$\text{أ} = \overline{\text{ص}} - \text{ب}\,\overline{\text{س}}$$

- التقدير يتم بالتعويض لقيمة س في معادلة خط الانحدار.
- مقدار الخطأ = |القيمة الجدولية − القيمة من معادلة الانحدار| = $|\text{ص}_{\text{س}} - \hat{\text{ص}}_{\text{س}}|$

الوحدة الثالثة

السلاسل الزمنيّة

Time Series

مشروع الوحدة: المياه واستهلاكها

١ **مقدمة المشروع:** تعتبر المياه وطريقة استهلاكها من أهم المشاكل في دولة الكويت وأكثرها تعقيدًا، نظرًا لمحدوديّة مواردها والمصادر المتجددة، ونظرًا لارتفاع معدلات استهلاكها مع مرور الوقت.

٢ **الهدف:** تحديد مصادر المياه ومحاولة توقع الكميات المطلوبة خلال الـ ٢٠ سنة القادمة بناء على عدة عوامل.

٣ **اللوازم:** شبكة الإنترنت، ورق رسم بياني، حاسوب.

٤ **أسئلة حول التطبيق:**

أ كيف كانت تؤمن دولة الكويت حاجاتها من المياه قبل تدفق عائدات النفط؟

ب ما كلفة إنتاج المياه العذبة المقطرة المحلاة؟ قارنها بكلفة الإنتاج في السنوات الماضية أي منذ ستينيات القرن الماضي. ارسم المضلع التكراري لكلفة تحلية المتر المكعب الواحد خلال الخمسين سنة الماضية آخذين بالاعتبار معدل الكلفة كل ٥ سنوات.

جـ ما المعدل اليومي لاستهلاك الفرد من المياه خلال الخمسين سنة الماضية. ارسم مضلعًا تكراريًّا يحدّد معدل الاستهلاك مع مرور الوقت آخذين بالاعتبار معدل الاستهلاك اليومي للفرد كل ٥ سنوات.

د قارن معدلات الاستهلاك بين عدة بلدان كقطر، والسعودية، وسلطنة عُمان في الفترات الزمنية ذاتها.

هـ ما معدل الزيادة السكانية في الكويت؟ وما تأثيره في السنوات القادمة على كمية المياه المستهلكة؟

٥ **التقرير:** قدّم تقريرًا مفصلًا عن هذا المشروع محاولًا توقع كميات الاستهلاك المطلوبة خلال الـ ٢٠ سنة القادمة، ومحددًا الموارد والمصادر التي يمكن اعتمادها لتأمين الحاجات مراعيًا الزيادة السكانية ليكون التقرير أكثر دقة وموضوعيّة.

دروس الوحدة

٣-١ السلسلة الزمنية	٣-٢ عناصر السلسلة الزمنية	٣-٣ تحليل السلاسل الزمنية
		معادلة الاتجاه العام للسلسلة الزمنية

الوحدة الثالثة

أين أنت الآن (المعارف السابقة المكتسبة)

- مخطط الانتشار.
- الارتباط وتطبيقاته.
- مُعامل ارتباط بيرسون.
- الانحدار وتطبيقاته.
- التقدير بمعادلة الانحدار.

ماذا سوف تتعلم؟

- السلسلة الزمنية.
- عناصر السلسلة الزمنية.
- تحليل السلاسل الزمنية.

المصطلحات الأساسية

السلسلة الزمنية – عناصر السلسلة الزمنية – المنحنى التاريخي للسلسلة الزمنية – الاتجاه العام – التغيرات الموسمية – التغيرات الدورية – التغيرات العرضية (الفجائية).

أضف إلى معلوماتك

تطور عمر الإنسان وزادت معدلاته، وذلك يعود إلى عدة عوامل أبرزها نوعية التغذية والرعاية الطبيّة، بحيث كان معدل عمر الإنسان عام ١٩٠٠ في الولايات المتحدة الأميركية حوالى ٤٧,٣ سنة ليصبح عام ٢٠٠٧ ٧٧,٩ سنة.

أما بالنسبة إلى الدول التي تعتبر فيها معدلات عمر الإنسان الأعلى في العالم، فتحل اليابان في المركز الثاني حيث إن معدل العمر فيها هو ٨٢ سنة، ودولة أندورة، التي تقع في جبال البيرينه بين فرنسا وإسبانيا، فتحتل في المركز الأول حيث يبلغ عدد سكانها ٧٢ ٠٠٠ نسمة ومعدل أعمار أبنائها ٨٣,٥ سنة. وبالتالي، فإن معدل عمر الإنسان في ارتفاع دائم مع مرور الزمن.

٣–١

السلسلة الزمنيّة

Time Series

سوف تتعلم
- السلسلة الزمنية.
- المنحنى التاريخي للسلسلة الزمنية.

دعنا نفكر ونتناقش

تعلّمت سابقًا كيف ترسم مخطط الانتشار لمتغيرين وكيفية إيجاد نوع العلاقة بينهما. في الجدول التالي:

س تمثّل السنوات ، ص تمثّل معدل النمو

س	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤	٢٠٠٥	٢٠٠٦	٢٠٠٧	٢٠٠٨	٢٠٠٩	٢٠١٠
ص	٢,٧	٢,٧	٢,٧	٢,٧	٢,٦	٢,٧	٢,٧	٢,٤	٢,٣	٢,٢	٢,١

أ مثّل البيانات بالخط المنكسر.

ب كيف كان معدّل النمو بين سنة ٢٠٠٠ وسنة ٢٠٠٦؟ وبعد سنة ٢٠٠٦؟

جـ ما نوع العلاقة بين الزمن ومعدّل النمو في هذه الفترات (ثابتة، متناقصة، متزايدة)؟

سبق لنا أن درسنا في الوحدة السابقة العلاقة بين ظاهرتين (متغيرين) من خلال موضوع الارتباط وفي هذه الوحدة سنتعرض لحالة خاصة من الارتباط بتثبيت قيم إحدى الظاهرتين (المتغيرين) وهو الزمن باعتباره المتغيّر المستقل ودراسة قيم الظاهرة الأخرى عبر الزمن وهو ما يسمى **بالسلسلة الزمنية**.

تعريف: السلسلة الزمنية

هي مجموعة القيم التي تأخذها ظاهرة ما في فترات زمنية غالبًا ما تكون متساوية ومتعاقبة.

أي أنها علاقة تربط بين متغيرين أحدهما هو قيم الظاهرة المطلوب دراستها والآخر هو الزمن. أي أننا نتتبع سلوك الظاهرة في أزمنة متعاقبة (سنة – نصف سنة – ربع سنة – شهر – يوم ...) ويسمى التتبع لقيم الظاهرة خلال هذه الأزمنة **بالسلسلة الزمنية**.

السلسلة الزمنية تحتوي على متغيرين أحدهما هو الزمن (المتغير المستقل) وسوف نرمز له بالرمز (س)، والآخر هو قيمة الظاهرة (المتغيّر التابع) وسنرمز له بالرمز (ص).

وتقاس قيم هذه الظواهر بنفس الوحدات ونفس طريقة القياس حتى يمكن المقارنة بين قيم الظاهرة خلال فترة الدراسة. وبعض السلاسل الزمنية تكون تصاعدية بصورة مطردة، وفي هذا النوع تزداد قيم الظاهرة محل الدراسة بمرور الزمن مثل إنتاج تحلية المياه في دولة الكويت، وبعض السلاسل الزمنية تكون تنازلية حيث تكون قيم مشاهداتها تتناقص بمرور الزمن مثل عدد الإصابات بشلل الأطفال في السنوات الأخيرة، والبعض الآخر من السلاسل الزمنية لا تخضع لنظام ثابت فهي متذبذبة بين التصاعدية والتنازلية وتكون قيم الظاهرة موزعة بين الصعود والنزول مثل إنتاج المشروبات الغازية على مدار السنة.

سوف يتم تمثيل السلسلة الزمنية بيانيًّا **بخط منكسر** ويسمى **بالمنحنى التاريخي للسلسلة** الزمنية ، حيث يتم تمثيل الزمن على المحور الأفقي والظاهرة على المحور الرأسي.

مثال (١)

يبيّن الجدول التالي متوسط العمر (ص) في إحدى الدول خلال السنوات (س) من سنة ٢٠٠٤ إلى سنة ٢٠١١.

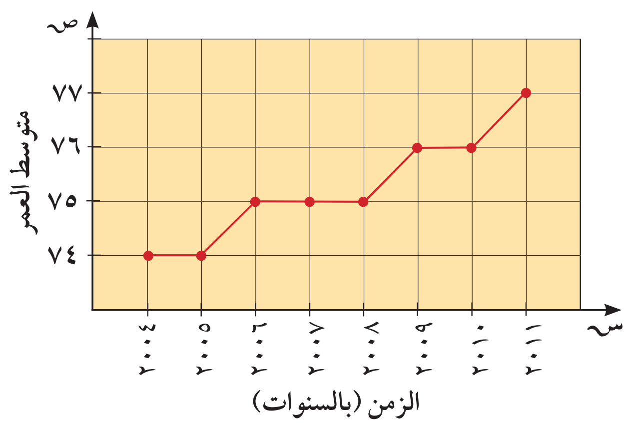
الزمن (س)	٢٠٠٤	٢٠٠٥	٢٠٠٦	٢٠٠٧	٢٠٠٨	٢٠٠٩	٢٠١٠	٢٠١١
العمر (ص)	٧٤	٧٤	٧٥	٧٥	٧٥	٧٦	٧٦	٧٧

أ مثّل بيانيًّا السلسلة الزمنية للبيانات الموجودة في الجدول أعلاه.

ب ما نوع العلاقة بين متوسط العمر والزمن؟

الحل:

أ مثّل الزمن على المحور الأفقي، ومتوسّط العمر على المحور الرأسي.



ب نلاحظ أن متوسط العمر في تزايد مع الزمن.

حاول أن تحل

١ في الجدول التالي متغيرين: الزمن (س) بالسنوات ، وعدد الولادات (ص) بالآلاف.

الزمن (س)	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤	٢٠٠٥	٢٠٠٦	٢٠٠٧	٢٠٠٨
عدد الولادات بالآلاف (ص)	٤٢	٤٢	٤٣	٤٥	٤٧	٥١	٥٣	٥٥	٥٥

أ مثّل بيانيًّا السلسلة الزمنية للبيانات الموجودة في الجدول أعلاه.

ب ما نوع العلاقة بين عدد الولادات والزمن؟

مثال (٢)

يبيّن الجدول التالي عدد الإصابات بشلل الأطفال (ص) بالآلاف في إحدى الدول خلال السنوات (س) من سنة ١٩٦٠ إلى سنة ١٩٦٧

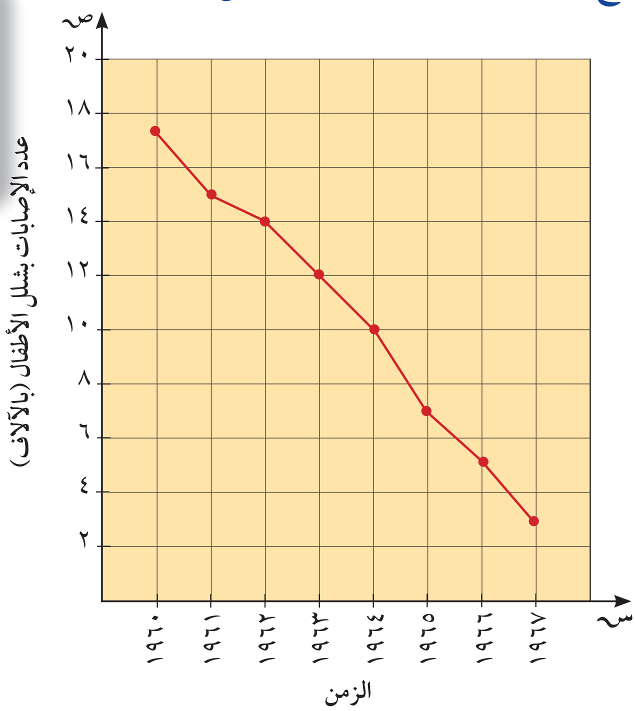
الزمن (س)	١٩٦٠	١٩٦١	١٩٦٢	١٩٦٣	١٩٦٤	١٩٦٥	١٩٦٦	١٩٦٧
عدد الإصابات بالآلاف (ص)	١٧	١٥	١٤	١٢	١٠	٧	٥	٣

أ مثّل بيانيًّا السلسلة الزمنية للبيانات الموجودة في الجدول أعلاه.

ب ما نوع العلاقة بين عدد الإصابات بشلل الأطفال والزمن؟

الحل:

أ



ب نلاحظ أن عدد الإصابات بشلل الأطفال في تناقص مع الزمن.

حاول أن تحل

٢ تهتم الدول بتنمية شعوبها من خلال القضاء على الأمّية باستخدام الحاسوب وذلك بإعداد برامج بهذا الخصوص، والجدول التالي يوضّح عدد الأميين بالمئات في محافظة ما من خلال الفترات الزمنية الموضّحة:

الزمن (س)	٢٠٠٢	٢٠٠٣	٢٠٠٤	٢٠٠٥	٢٠٠٦	٢٠٠٧	٢٠٠٨	٢٠٠٩	٢٠١٠
عدد الأميين بالمئات (ص)	٣١	٢٧	٢٥	٢٥	٢٤	٢٥	٢٣	٢١	١٩

أ مثّل بيانيًّا السلسلة الزمنية للبيانات الموجودة في الجدول أعلاه.

ب ما نوع العلاقة بين عدد الأميين في استخدام الحاسوب والزمن؟

عناصر السلسلة الزمنيّة

Time Series Elements

سوف تتعلم
- الاتجاه العام.
- التغيرات الموسمية.
- التغيرات الدورية.
- التغيرات العرضية.

دعنا نفكر ونتناقش

انظر إلى السلاسل الزمنية التالية:

١ قارن فيما بينها.

٢ اقترح أمثلة حياتية تتطابق مع السلسلة الزمنية في الشكل أ .

٣ أي سلسلة من السلاسل الزمنية الثلاث تبيّن تغيرًا فجائيًا؟

درسنا فيما سبق أن السلسلة الزمنية هي علاقة بين متغيّرين أحدهما يسمى المتغيّر المستقل وهو الزمن (س)، والآخر يسمى المتغيّر التابع (ص)، ويوجد عدد من المؤثرات المشتركة في كل سلسلة زمنية ولكنها تؤثر بدرجات مختلفة من ظاهرة لأخرى طبقًا لطبيعة الظاهرة محل الدراسة.

والهدف من الدراسة الإحصائية للسلسلة الزمنية هو اكتشاف التغيرات التي تطرأ على قيم الظاهرة من زيادة أو نقصان في زمن محدد وتسمى هذه التغيرات التي تؤثر على السلسلة الزمنية سواء كانت مجتمعة أم منفردة **بعناصر السلسلة الزمنية.**

عناصر السلسلة الزمنية هي:

١ المؤثرات الاتجاهية (الاتجاه العام للسلسلة الزمنية).

٢ التغيرات الموسمية.

٣ التغيرات الدورية.

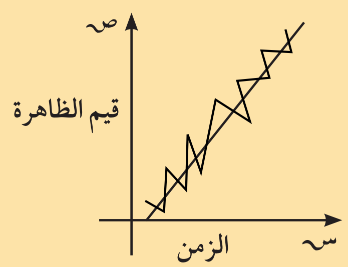
٤ التغيرات العرضية (الفجائية).

وسنتناول هذه العناصر بشيء من التفصيل.

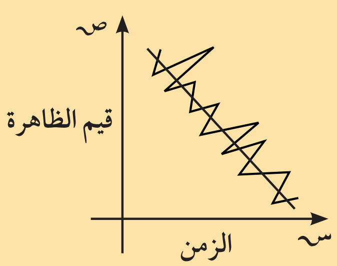
١ـ الاتجاه العام للسلسلة الزمنية Secular Trend

الاتجاه العام للسلسلة الزمنية هو الاتجاه الذي تأخذه السلسلة الزمنية لحدث ما خلال فترة طويلة من الزمن.

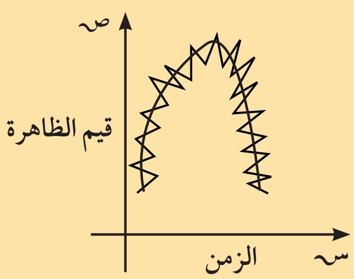
هناك العديد من الأمثلة التي تبيّن ذلك منها: عدد سكان بلد ما، الفئات العمريّة للمجتمع، ...



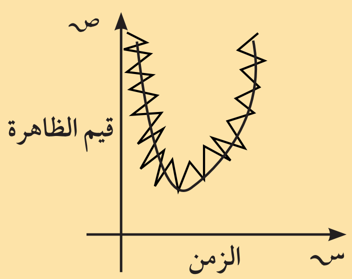
الاتجاه العام في زيادة مستمرة



الاتجاه العام في نقصان مستمر



الاتجاه العام في زيادة ثم نقصان

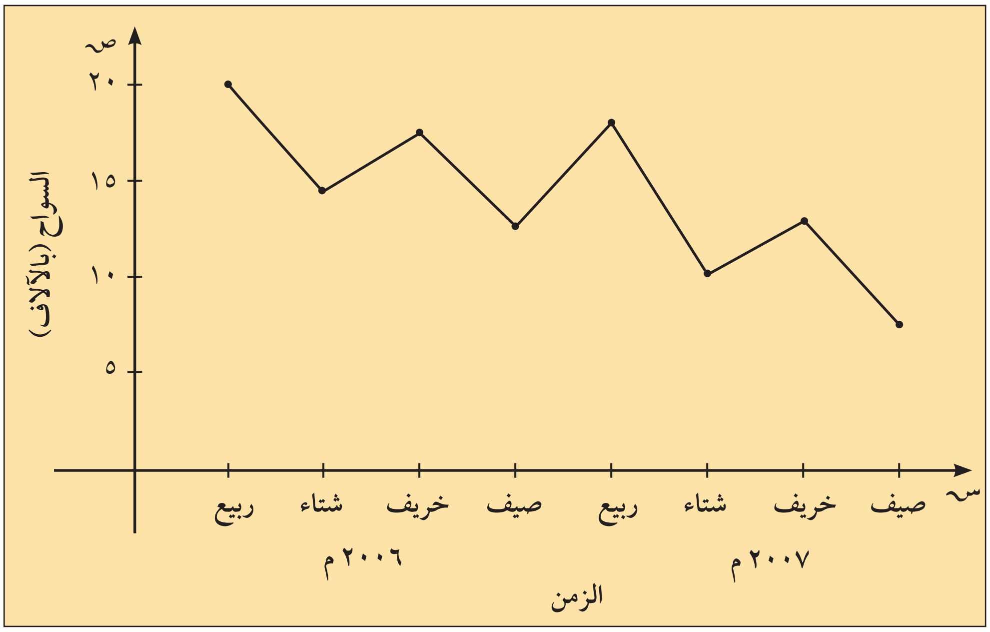


الاتجاه العام في نقصان ثم زيادة

٢ـ التغيرات الموسمية Seasonal Variations

هي التغيرات التي تتكرر بانتظام خلال فترات زمنية أقل من سنة كأن تكون نصف سنوية أو ربع سنوية أو شهرية أو أسبوعية أو. . .

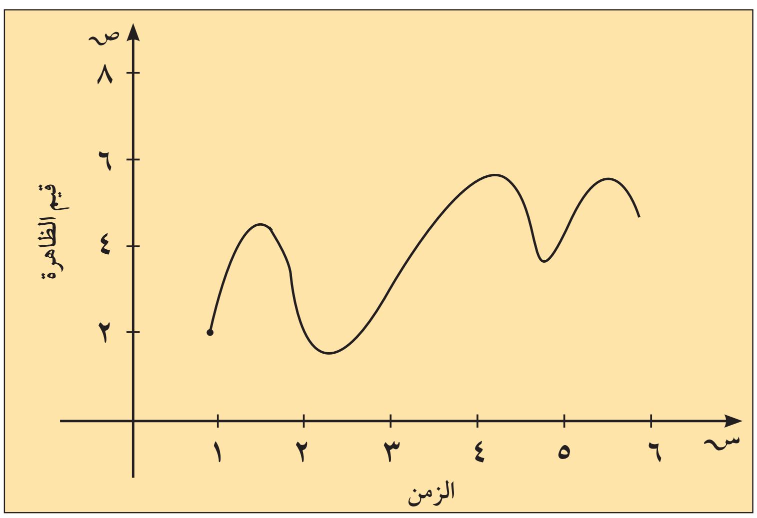
والأمثلة على ذلك متعددة منها سقوط الأمطار بشكل موسمي، وكذلك مبيعات المشروبات الغازية تزداد خلال فصل الصيف، واستهلاك الكهرباء والماء يزداد أيضًا في فصل الصيف، وزيادة حركة المواصلات وازدحام الطرق في فترتي الصباح والظهيرة من كل يوم، والشكل التالي يبيّن التغيرات الموسمية لأعداد السواح بالآلاف للعامين ٢٠٠٦ م، ٢٠٠٧ م على الترتيب.



لاحظ أن الاتجاه العام للسلسلة الزمنية في نقصان.

٣– التغيرات الدورية Cyclic Variations

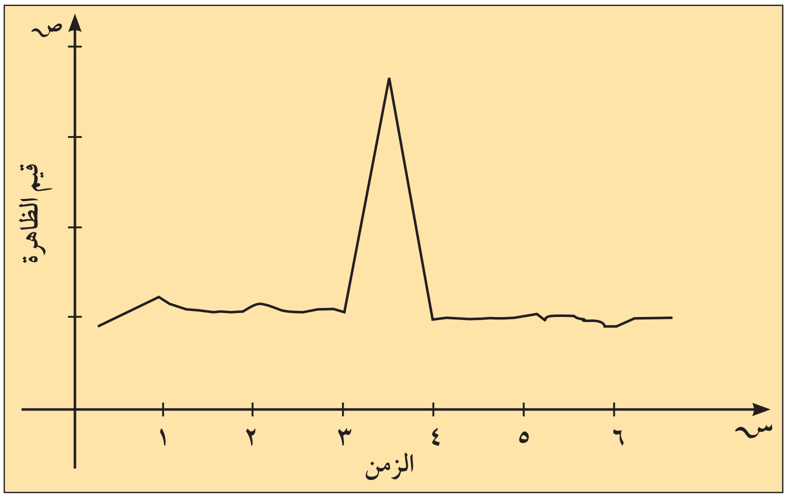
هي تغيرات للسلسلة الزمنية على فترات طويلة المدى نسبيًّا أكثر من سنة، وتختلف التغيرات الدورية عن التغيرات الموسمية في أن التغيرات الموسمية تحدث في فترات زمنية أقل من سنة، ويمكن اعتبار التغيرات الدورية تحركًا لفترة أقل طولًا من فترة الاتجاه العام، ومن الأمثلة المهمة للتغيرات الدورية ما يحدث لشركة ما من فترة رخاء اقتصادي، ثم فترة ركود اقتصادي، ثم فترة كساد، ثم انفراج من الأزمة الاقتصادية كما هو موضح في الشكل.



لاحظ أن الاتجاه العام للسلسلة في تزايد.

٤ – التغيرات العرضية (الفجائية) Irregular Variations

تتأثر كثير من الظواهر من وقت إلى آخر بعوامل مختلفة تعود إلى تغيرات **غير متوقعة** أو إلى **أمور يصعب التنبؤ بها**، فمثلًا في المحلات التجارية تختلف قيم المبيعات من يوم إلى آخر متأثرة بطبيعة الطقس أو وجود حفلات زواج وما إلى ذلك من تغيرات. كما أن التغيرات تحدث نتيجة **عوامل مفاجئة** كالحروب، والفيضانات، والأوبئة، والزلازل. والتغيرات من هذا النوع تعرف بالتغيرات العرضية أو الفجائية، ويمكن توضيح التغيرات العرضية أو الفجائية في المنحنى التاريخي للسلسلة الزمنية بالشكل التالي:



مثال (١)

يمثّل الجدول التالي أرباح إحدى الشركات الكبرى بملايين الدنانير من سنة ١٩٨٥ إلى سنة ٢٠٠٠

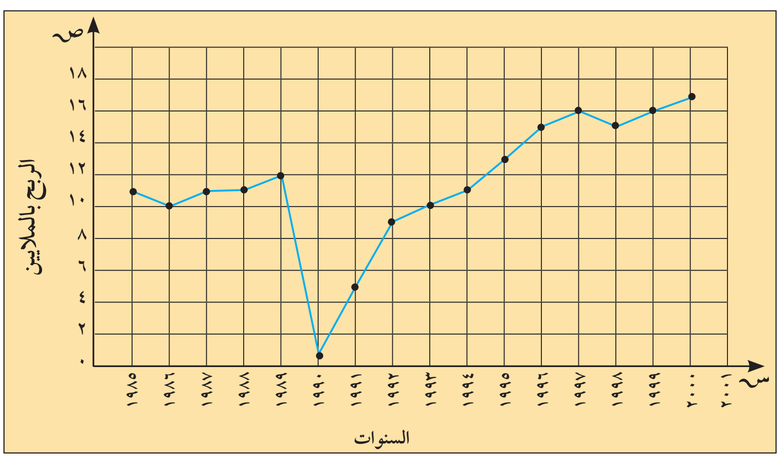
السنة (س)	٨٥	٨٦	٨٧	٨٨	٨٩	٩٠	٩١	٩٢	٩٣	٩٤	٩٥	٩٦	٩٧	٩٨	٩٩	٢٠٠٠
الربح بالملايين (ص)	١١	١٠	١١	١١	١٢	١	٥	٩	١٠	١١	١٣	١٥	١٦	١٥	١٦	١٧

أ مثّل بيانيًّا على شكل خط منكسر بيانات الجدول أعلاه.

ب ما نوع التغيّرات التي طرأت على أرباح هذه الشركة؟ وما السبب الأبرز لهذه التغيّرات؟

الحل:

أ



ب لدينا تغيّر مفاجئ في سنة ١٩٩٠ يمثّل بانخفاض جذري للأرباح.
السبب الأبرز هو العدوان العراقي على الكويت.

حاول أن تحل

١ يبيّن الجدول التالي عدد المنتسبين إلى أحد الأندية الرياضية خلال أشهر سنة ٢٠٠٨

الأشهر (س)	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢
عدد المنتسبين (ص)	٣٠	٣٢	٤٠	٤١	٥٠	٥٠	٦٠	٧٠	٧٥	٧١	٦٠	٥٥

أ مثّل بيانيًا على شكل خط منكسر بيانات الجدول أعلاه.

ب ما الذي تلاحظه في الرسم البياني؟

جـ برأيك، ما سبّب هذه التغيّرات؟

مثال (٢)

يبيّن الجدول التالي عدد البشوت المباعة في أحد المجمعات التجارية خلال فترة زمنية من أربعة أشهر وعلى امتداد أربع سنوات.

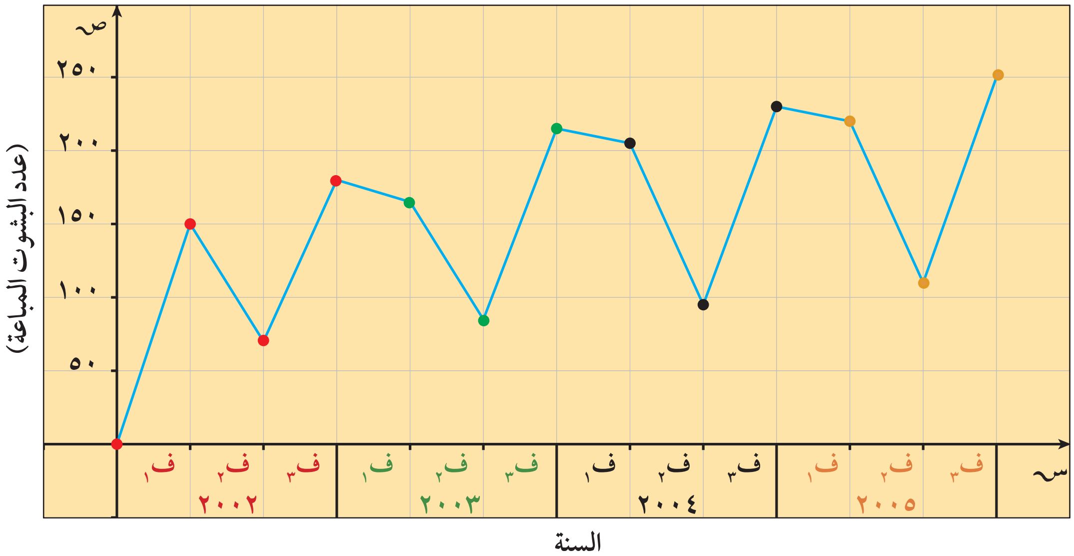
الفترة / السنوات	الأولى	الثانية	الثالثة
٢٠٠٢	١٥٠	٧٠	١٨٠
٢٠٠٣	١٦٥	٨٥	٢١٥
٢٠٠٤	٢٠٥	٩٥	٢٣٠
٢٠٠٥	٢٢٠	١١٠	٢٥٠

أ مثّل بيانيًا على شكل خط منكسر بيانات الجدول أعلاه.

ب ما الذي تلاحظه؟

الحل:

أ



ب تتكرر التغيرات بانتظام خلال الفترات الزمنية من ٤ أشهر. تزداد المبيعات في الفترتين الأولى والثالثة من كل سنة مع ازدياد خفيف خلال السنوات.

حاول أن تحل

٢ يبيّن الجدول التالي مبيعات إحدى المؤسسات التجارية (بآلاف الدنانير) خلال كل فصل من فصول السنة الأربعة وعلى امتداد ثلاث سنوات.

الفصل / السنة	الأول	الثاني	الثالث	الرابع
٢٠٠٣	٢٠٢	١٥٠	٥٠	١٠٠
٢٠٠٤	٢١٠	١٧٠	٦٠	١١٠
٢٠٠٥	٢٣٠	١٩٠	٧٥	١٣٠

أ مثّل بيانيًّا على شكل خط منكسر بيانات الجدول.

ب ما الذي تلاحظه؟

مثال (٣)

يبيّن الجدول التالي مبيعات إحدى الشركات (بمئات آلاف الدنانير) خلال فترة ثماني سنوات موزعة على كل نصف سنة كما في الجدول التالي:

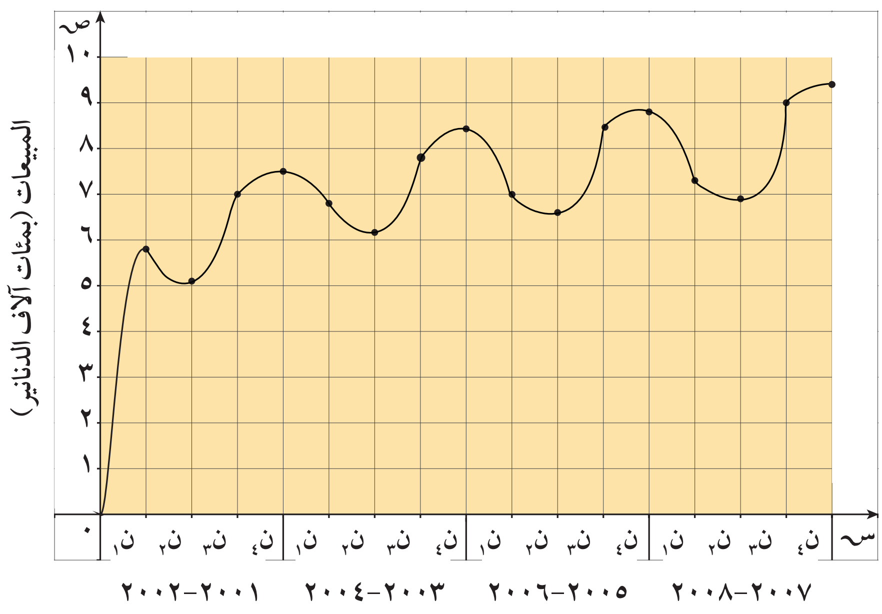
نصف السنة / السنوات	النصف الأول	النصف الثاني	النصف الثالث	النصف الرابع
٢٠٠١ – ٢٠٠٢	٥,٨	٥,١	٧,٠	٧,٥
٢٠٠٣ – ٢٠٠٤	٦,٨	٦,٢	٧,٨	٨,٤
٢٠٠٥ – ٢٠٠٦	٧,٠	٦,٦	٨,٥	٨,٨
٢٠٠٧ – ٢٠٠٨	٧,٣	٦,٩	٩,٠	٩,٤

أ ارسم بيانيًّا على شكل منحنى بيانات الجدول أعلاه.

ب ما الذي تلاحظه بالنسبة إلى الاتجاه العام للسلسلة؟

الحل:

أ



ب الاتجاه العام للسلسلة في تزايد.

حاول أن تحل

٣ يبيّن الجدول التالي المسافات التي يركضها (بعشرات الأمتار) أحد لاعبي كرة القدم خلال ١٤ دقيقة.

الزمن	١	٢	٣	٤	٥	٦	٧	٨	٩	١٠	١١	١٢	١٣	١٤
المسافة (بعشرات الأمتار)	١٥	٧	٩	٨	٢	٦	١٤	١٥	٧	٩	٨	٢	٦	١٤

أ ارسم بيانيًّا على شكل منحنى بيانات الجدول أعلاه.

ب ما الذي تلاحظه بالنسبة إلى الاتجاه العام للسلسلة؟

٣ـــ٣

تحليل السلاسل الزمنية

Analysing Time Series

سوف تتعلم

- معادلة الاتجاه العام للسلسلة الزمنية.
- حساب مقدار الخطأ

دعنا نفكر ونتناقش

أخذت أوزان عشرة أطفال عند الولادة في أحد المستشفيات الغربية بهدف دراسة العلاقة بين وزن الطفل عند الولادة وعدد السجائر التي تدخنها الأم يوميًّا خلال أول شهرين من فترة الحمل.

عدد السجائر في اليوم (س)	١٥	١٠	٥	٨	٩	٧	١١	٦	٣	٢
الوزن بالجرام (ص)	١٤٤٧	١٥٠٠	١٧٠١	١٧١٢	١٨٥٧	٢٠٣١	٢١٤٥	٢٢١٤	٢٢١٠	٢٥٣٧

أ هل يوجد علاقة بين المتغيرين س ، ص؟

(إرشاد: أوجد مُعامل الارتباط (ر))

ب أوجد معادلة خط الانحدار.

جـ إذا كان وزن الطفل عند الولادة ١٩٥٠ جرامًا،

فما تقريبًا عدد السجائر التي تدخنها الأم يوميًّا؟

معادلة الاتجاه العام للسلسلة الزمنية Equation of Time Series

الاتجاه العام للسلسلة الزمنية هو أهم عنصر من عناصر السلسلة ، لأنه يساعد الباحثين وذوي الاختصاص على تقدير أو توقع قيمة مستقبلية لزمن قادم.

تعلّمنا سابقًا كيفية إيجاد معادلة خطّ الانحدار.

وفي هذا الدرس ، سوف نستخدم الطريقة ذاتها لإيجاد معادلة الاتجاه العام للسلسلة الزمنية مع فرق بسيط وهو استخدام المتغيّر (س) لتمثيل الزمن، بفرض أن العلاقة بين الزمن (س) وقيم الظاهرة (ص) هي علاقة خطيّة.

الخطوات المتبعة لإيجاد معادلة الاتجاه العام للسلسلة الزمنية

١ نفرض قيم الزمن (س) باعتباره الفترة الأولى (سنة الأساس) ونعبر عنه بالعدد صفر، الفترة الثانية بالعدد ١، ثم الفترة الثالثة بالعدد ٢، وهكذا ...

٢ نعيّن قيم الثوابت أ ، ب كما سبق شرحه حيث:

$$ب = \frac{ن(\sum س ص) - (\sum س)(\sum ص)}{ن(\sum س^{٢}) - (\sum س)^{٢}}$$

$$أ = \overline{ص} - ب\,\overline{س} \qquad حيث: \overline{ص} = \frac{\sum ص}{ن} ، \overline{س} = \frac{\sum س}{ن}$$

٣ معادلة الاتجاه العام تكتب على الشكل التالي: $\hat{ص} = أ + ب\,س$

٤ يمكننا التنبؤ بقيمة ص إذا علمت قيمة س.

٥ نحسب مقدار الخطأ:

مقدار الخطأ = |القيمة الجدولية – القيمة التي تحقق معادلة الاتجاه العام للسلسلة الزمنية|

ونعبر عنه بـ : $|ص_{س} - \hat{ص}_{س}|$.

مثال (١)

يبيّن الجدول التالي عدد الخبراء الأجانب بالآلاف في دولة ما، من سنة ٢٠٠٧ حتى سنة ٢٠١٤

السنوات (س)	٢٠٠٧	٢٠٠٨	٢٠٠٩	٢٠١٠	٢٠١١	٢٠١٢	٢٠١٣	٢٠١٤
عدد الخبراء بالآلاف (ص)	٠,٥	٠,٧	٠,٨٣	١,٢	١,٥	١,٨	١,٣	١

أ أوجد معادلة الاتجاه العام لعدد الخبراء الأجانب في الفترة المذكورة أعلاه.

ب قدّر كم سيصبح عدد الخبراء سنة ٢٠١٧

جـ احسب مقدار الخطأ في عدد الخبراء سنة ٢٠١٢

الحل:

أ نعتبر سنة ٢٠٠٧ هي السنة الأساس ونعبّر عنها بالعدد صفر، وسنة ٢٠٠٨ بالعدد ١

وهكذا دواليك حتى سنة ٢٠١٤ فنعبّر عنها بالعدد ٧

السنوات	س	ص	س ص	$س^٢$
٢٠٠٧	٠	٠,٥	٠	٠
٢٠٠٨	١	٠,٧	٠,٧	١
٢٠٠٩	٢	٠,٨٣	١,٦٦	٤
٢٠١٠	٣	١,٢	٣,٦	٩
٢٠١١	٤	١,٥	٦	١٦
٢٠١٢	٥	١,٨	٩	٢٥
٢٠١٣	٦	١,٣	٧,٨	٣٦
٢٠١٤	٧	١	٧	٤٩
المجموع	$\sum$ س = ٢٨	$\sum$ ص = ٨,٨٣	$\sum$ س ص = ٣٥,٧٦	$\sum$ $س^٢$ = ١٤٠

$$ب = \frac{ن(\sum س ص) - (\sum س)(\sum ص)}{ن(\sum س^٢) - (\sum س)^٢}$$

$$ب = \frac{٨(٣٥,٧٦) - (٢٨)(٨,٨٣)}{٨(١٤٠) - ٧٨٤}$$

$$ب \approx ٠,١١٥٦$$

$$أ = \overline{ص} - ب\,\overline{س}$$

$$\overline{ص} = \frac{\sum ص}{ن} = ١,١٠٣٨ \quad ، \quad \overline{س} = \frac{\sum س}{ن} = ٣,٥$$

$$أ = ١,١٠٣٨ - ٠,١١٥٦ \times ٣,٥$$

$$أ = ٠,٦٩٩٢$$

$\therefore$ معادلة الاتجاه العام هي:

$$\hat{ص} = أ + ب\,س$$

$$\hat{ص} = ٠,٦٩٩٢ + ٠,١١٥٦\,س$$

ب نريد تقدير عدد الخبراء الأجانب سنة ٢٠١٧، أي عند س = ١٠

$$\hat{ص} = ٠,٦٩٩٢ + ٠,١١٥٦ \times ١٠$$

$$\hat{ص} = ١,٨٥٥٢$$

تقدير سنة ٢٠١٧ هو ١٨٥٥ خبيرًا أجنبيًّا (١,٨٥٥٢ × ١٠٠٠ = ١٨٥٥,٢)

ج $ص̂ = أ + ب س$

$ص̂_{٢٠١٢} = ٠,٦٩٩٢ + ٠,١١٥٦ × ٥ = ١,٢٧٧٢$

مقدار الخطأ = $|١,٨ - ١,٢٧٧٢| = ٠,٥٢٢٨$

أي أن مقدار الخطأ في عدد الخبراء $٠,٥٢٢٨ × ١٠٠٠ = ٥٢٢,٨ ≈ ٥٢٣$ خبيرًا

حاول أن تحل

١ يبيّن الجدول التالي عدد مستخدمي شبكة الإنترنت بالآلاف في دولة ما من سنة ٢٠٠٠ حتى سنة ٢٠٠٨

السنوات (س)	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤	٢٠٠٥	٢٠٠٦	٢٠٠٧	٢٠٠٨
عدد المستخدمين (بالآلاف) (ص)	١٠٠	١٥٠	٢٠٠	٧٦٧	٦٣٣	٧٠٠	٨٠٠	٩٠٠	١٠٠٠

أ أوجد معادلة الاتجاه العام.

ب قدّر عدد مستخدمي شبكة الإنترنت سنة ٢٠١٢

ج أوجد مقدار الخطأ سنة ٢٠٠٦

مثال (٢)

يبيّن الجدول التالي التكلفة لإنتاج إحدى السلع بالألف دينار كويتي من سنة ٢٠٠٦ حتى سنة ٢٠١٣

السنة (س)	٢٠٠٦	٢٠٠٧	٢٠٠٨	٢٠٠٩	٢٠١٠	٢٠١١	٢٠١٢	٢٠١٣
التكلفة (بالألف دينار)(ص)	١٥	١٦	١٨	١٨	٢٠	٢٢	٢٤	٢٨

أ أوجد معادلة الاتجاه العام لتكلفة إنتاج السلعة.

ب قدّر قيمة التكلفة عام ٢٠١٧

ج احسب مقدار الخطأ سنة ٢٠١١

الحل:

أ نعتبر سنة ٢٠٠٦ هي السنة الأساس.

السنوات	س	ص	س ص	$س^2$
٢٠٠٦	٠	١٥	٠	٠
٢٠٠٧	١	١٦	١٦	١
٢٠٠٨	٢	١٨	٣٦	٤
٢٠٠٩	٣	١٨	٥٤	٩
٢٠١٠	٤	٢٠	٨٠	١٦
٢٠١١	٥	٢٢	١١٠	٢٥
٢٠١٢	٦	٢٤	١٤٤	٣٦
٢٠١٣	٧	٢٨	١٩٦	٤٩
المجموع	$\sum$ س = ٢٨	$\sum$ ص = ١٦١	$\sum$ س ص = ٦٣٦	$\sum س^2$ = ١٤٠

$$ن = ٨ ، \overline{س} = \frac{\sum س}{ن} = \frac{٢٨}{٨} = ٣,٥ ، \overline{ص} = \frac{\sum ص}{ن} = \frac{١٦١}{٨} = ٢٠,١٢٥$$

$$ب = \frac{ن(\sum س ص) - (\sum س)(\sum ص)}{ن(\sum س^2) - (\sum س)^2} = \frac{٨ \times ٦٣٦ - ٢٨ \times ١٦١}{٨ \times ١٤٠ - (٢٨)^2}$$

$$= \frac{٥٨٠}{٣٣٦} \approx ١,٧٢٦٢$$

$$أ = \overline{ص} - ب \overline{س} \leftarrow أ = ٢٠,١٢٥ - (١,٧٢٦٢) \times ٣,٥ = ١٤,٠٨٣٣$$

$\therefore$ معادلة الاتجاه العام هي:

$$\hat{ص} = ١٤,٠٨٣٣ + ١,٧٢٦٢ \text{ س}$$

ب قيمة التكلفة سنة ٢٠١٧ عند س = ١١

$$\therefore \hat{ص}_{٢٠١٧} = ١٤,٠٨٣٣ + ١,٧٢٦٢ \times ١١$$

= ٣٣,٠٧١٥ ألف دينار

جـ سنة ٢٠١١ ← $ص_{٢٠١١} = ٢٢$

$$\hat{ص}_{٢٠١١} = ١٤,٠٨٣٣ + ١,٧٢٦٢ \times ٥$$

$$= ٢٢,٧١٤٣$$

∴ مقدار الخطأ = |ص$_{٢٠١١}$ − صَ$_{٢٠١١}$|

= |٢٢ − ٢٢,٧١٤٣|

= ٠,٧١٤٣

∴ مقدار الخطأ = ٧١٤,٣ ديناراً

حاول أن تحل

٢ الجدول التالي يبيّن قيم ظاهرة معينة خلال ٧ سنوات.

السنة	١٩٩٨	١٩٩٩	٢٠٠٠	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤
قيم الظاهرة	٣	٥	٨	١٠	١٤	١٦	١٨

أ أوجد معادلة الاتجاه العام لقيم الظاهرة.

ب تنبأ بالقيمة المتوقعة للظاهرة سنة ٢٠٠٧

جـ احسب مقدار الخطأ سنة ٢٠٠٣

مثال (٣)

الجدول التالي يبيّن إنتاج إحدى شركات السيارات بالألف سيارة من سنة ٢٠٠٧ حتى سنة ٢٠١٣

السنة (س)	٢٠٠٧	٢٠٠٨	٢٠٠٩	٢٠١٠	٢٠١١	٢٠١٢	٢٠١٣
عدد السيارات بالألف (ص)	٤٠	٦٠	٧٠	٩٠	١٠٠	١٥٠	١٨٠

أ أوجد معادلة الاتجاه العام للسلسلة الزمنية

ب قدّر عدد السيارات المنتجة سنة ٢٠١٦

جـ احسب مقدار الخطأ سنة ٢٠١١

الحل:

أ نعتبر سنة ٢٠٠٧ هي السنة الأساس.

السنوات	س	ص	س ص	$\text{س}^{٢}$
٢٠٠٧	٠	٤٠	٠	٠
٢٠٠٨	١	٦٠	٦٠	١
٢٠٠٩	٢	٧٠	١٤٠	٤
٢٠١٠	٣	٩٠	٢٧٠	٩
٢٠١١	٤	١٠٠	٤٠٠	١٦
٢٠١٢	٥	١٥٠	٧٥٠	٢٥
٢٠١٣	٦	١٨٠	١٠٨٠	٣٦
المجموع	٢١	٦٩٠	٢٧٠٠	٩١

$\text{ن} = ٧$ ، $\overline{\text{س}} = \frac{\sum \text{س}}{\text{ن}} = \frac{٢١}{٧} = ٣$ ، $\overline{\text{ص}} = \frac{\sum \text{ص}}{\text{ن}} = \frac{٦٩٠}{٧} \approx ٩٨,٥٧١٤$

$$\text{ب} = \frac{\text{ن}(\sum \text{س ص}) - (\sum \text{س})(\sum \text{ص})}{\text{ن}(\sum \text{س}^{٢}) - (\sum \text{س})^{٢}} = \frac{٧ \times ٢٧٠٠ - ٢١ \times ٦٩٠}{٧ \times ٩١ - (٢١)^{٢}}$$

$= ٢٢,٥$

$\because$ $\text{أ} = \overline{\text{ص}} - \text{ب}\,\overline{\text{س}} \leftarrow \text{أ} = ٩٨,٥٧١٤ - ٢٢,٥ \times ٣$

$= ٣١,٠٧١٤$

$\because$ معادلة الاتجاه العام هي: $\hat{\text{ص}} = \text{أ} + \text{ب س}$

$\hat{\text{ص}} = ٣١,٠٧١٤ + ٢٢,٥ \text{ س}$

ب تقدير عدد السيارات المنتجة سنة ٢٠١٦ أي عند س = ٩

$\hat{\text{ص}} = ٣١,٠٧١٤ + ٢٢,٥ \times ٩$

$\hat{\text{ص}} = ٢٣٣,٥٧١٤$

تقدير عدد السيارات المنتجة سنة ٢٠١٦ هو حوالى ٢٣٤ ألف سيارة.

ج $\text{ص}_{٢٠١١} = ١٠٠$ ، $\hat{\text{ص}}_{٢٠١١} = ٣١,٠٧١٤ + ٢٢,٥ \times ٤$

$= ١٢١,٠٧١٤$

مقدار الخطأ = $|\text{ص}_{٢٠١١} - \hat{\text{ص}}_{٢٠١١}| = |١٠٠ - ١٢١,٠٧١٤|$

$= ٢١,٠٧١٤$

حوالى ٢١ ألف سيارة.

حاول أن تحل

٣ الجدول التالي يوضّح مبيعات إحدى الشركات بالألف دينار في الفترة من سنة ٢٠٠١ وحتى سنة ٢٠٠٧

السنة (س)	٢٠٠١	٢٠٠٢	٢٠٠٣	٢٠٠٤	٢٠٠٥	٢٠٠٦	٢٠٠٧
المبيعات بالألف (ص)	٨٧	٩١	٩٦	١٠٩	١١٩	١٢٩	١٣٥

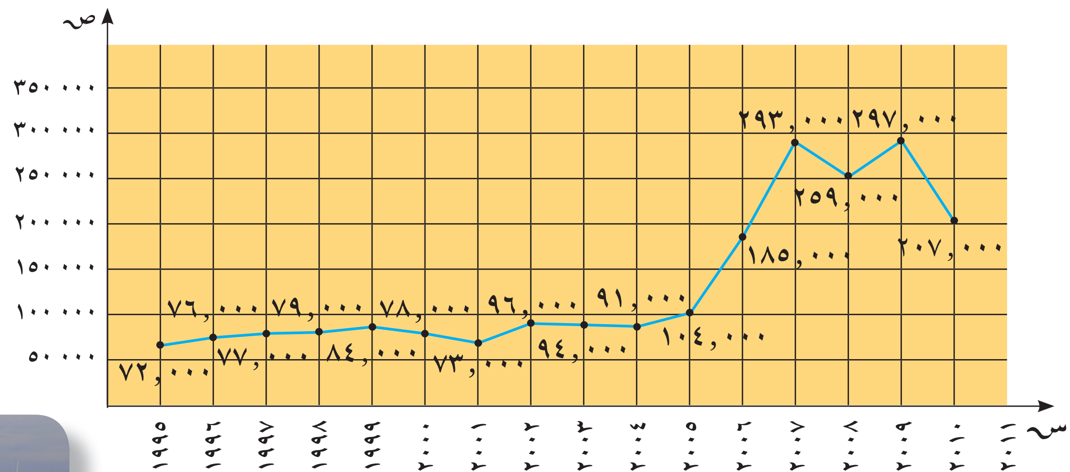
أوجد:

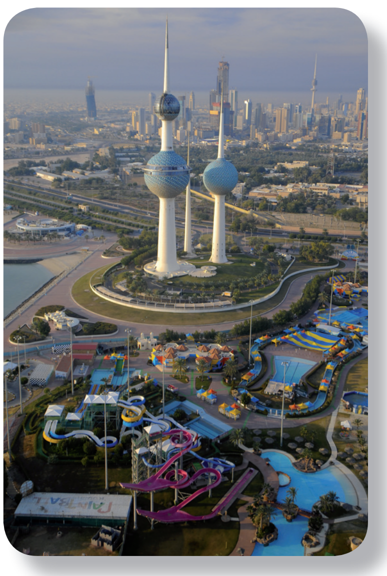
أ معادلة خط الاتجاه العام للمبيعات خلال الفترة المذكورة.

ب القيمة المتوقعة للمبيعات عام ٢٠١٠

جـ مقدار الخطأ سنة ٢٠٠٥

يبيّن الخط المنكسر التالي أعداد السواح الذين قاموا بزيارة دولة الكويت من سنة ١٩٩٥ حتى سنة ٢٠١٠





أ كوّن جدولًا مستخدمًا المعطيات من الرسم البياني للخط المنكسر.

ب أوجد معادلة الاتجاه العام.

ج قدّر عدد السواح لسنة ٢٠١٥

د أوجد مقدار الخطأ سنة ٢٠١٠

الحل:

يهتم المعنيّون بتقديرات عدد السواح للأعوام القادمة، وبإيجاد مقدار الخطأ لسنة ٢٠١٠

أ نستخرج المعلومات من الخط المنكسر ونضعها في جدول على الشكل التالي:

السنوات	س	ص	س ص	$س^٢$
١٩٩٥	٠	٧٢ ٠٠٠	٠	٠
١٩٩٦	١	٧٦ ٠٠٠	٧٦ ٠٠٠	١
⋮	⋮	⋮	⋮	⋮
٢٠١٠	١٥	٢٠٧ ٠٠٠	٣١٠ ٥٠٠	٢٢٥
	$\sum$ س = ١٢٠	$\sum$ ص = ٢ ١٦٥ ٠٠٠	$\sum$ س ص = ٢١ ٠٩٥ ٠٠٠	$\sum س^٢$ = ١ ٢٤٠

معادلة الاتجاه العام:

ب أ = ٢٨ ١٦١,٨ ، ب = ١٤ ٢٨٦,٨ ص̂ = ٢٨ ١٦١,٨ + ١٤ ٢٨٦,٨س

جـ نقدّر سنة ٢٠١٥ ، س = ٢٠ ، بالتعويض بـ «$\hat{ص}$»:

$$\hat{ص}_{٢٠١٥} \approx ٣١٣٨٩٧$$

د نوجد مقدار الخطأ لسنة ٢٠١٠:

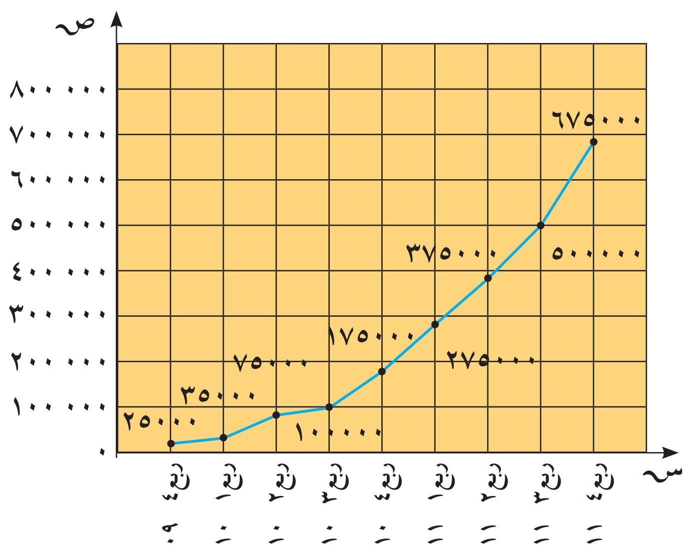
$$\text{مقدار الخطأ} = |ص_{٢٠١٠} - \hat{ص}_{٢٠١٠}| = ٣٥\ ٤٦٣,٨$$

$$\approx ٣٥\ ٤٦٤$$

مقدار الخطأ تقريبًا ٣٥ ٤٦٤ سائحًا.

مسألة إضافية

يمثّل الخط المنكسر التالي تطور عدد تطبيقات الهواتف الذكية التي تعمل بحسب أحد أنظمة التشغيل وذلك خلال الأرباع التالية من الربع الرابع من سنة ٢٠٠٩ إلى الربع الرابع من سنة ٢٠١١



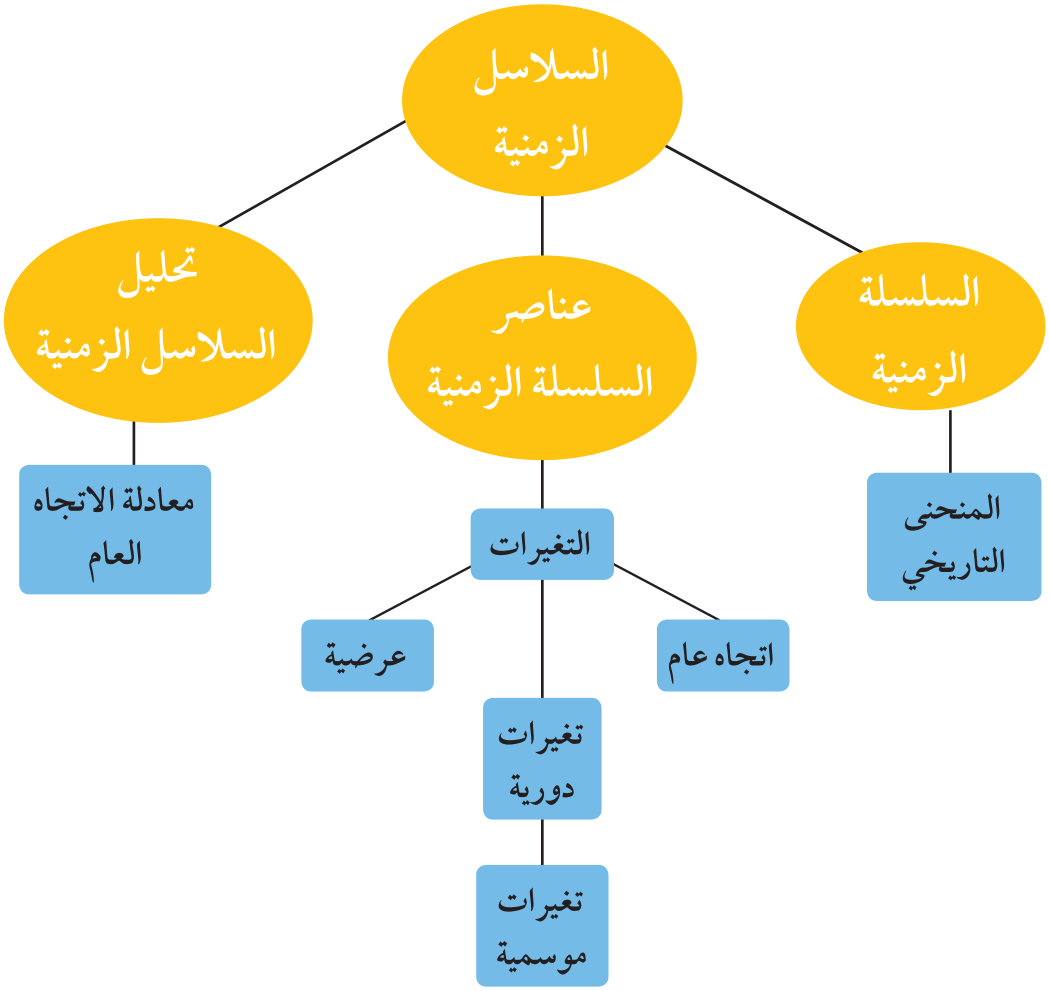
يهتم المعنيّون بمعرفة تطور أعداد التطبيقات في الربع الرابع من سنة ٢٠١٥ لما يترتب على ذلك من ارتفاع في المداخيل من جراء تحميل هذه التطبيقات في الهواتف الذكية.

أ كوّن جدولًا كما في «المرشد لحل المسائل» مستخرجًا المعطيات من الرسم البياني للخط المنكسر.

ب ما هو العدد المتوقع للتطبيقات في الربع الرابع من سنة ٢٠١٥؟

جـ ما هو مقدار الخطأ في الربع الرابع من سنة ٢٠١٠؟

مخطط تنظيمي للوحدة الثالثة



ملخص

- السلسلة الزمنية هي مجموعة قيم تأخذها ظاهرة ما في فترات زمنيّة مختلفة.
- المنحنى التاريخي للسلسلة الزمنية هو الخط المنكسر الذي يربط النقاط الممثلة للبيانات.
- الاتجاه العام هو الاتجاه الذي تأخذه السلسلة على مدّة طويلة من الزمن.
- الاتجاه العام للسلسلة يمكن أن يكون تصاعديًّا أو تنازليًّا أو كليهما معًا.
- التغيرات الموسمية هي تغيرات تتكرّر بانتظام خلال فترات معيّنة من الزمن تكون مدّتها أقل من سنة.
- التغيرات الدورية هي تغيرات على فترة طويلة المدى أي أكثر من سنة.
- التغيرات العرضية هي تغيرات فجائية تعود إلى الصدفة البحتة أو إلى أمور يصعب تكهنها.
- معادلة الاتجاه العام تستخدم في عملية التكهن بقيم الظاهرة لفترات زمنية مستقبليّة. وتعطى بالقاعدة:

$$\hat{ص} = أ + ب\,س$$

حيث: $ب = \dfrac{ن \left(\sum س\,ص\right) - \left(\sum س\right)\left(\sum ص\right)}{ن\left(\sum س^{٢}\right) - \left(\sum س\right)^{٢}}$ ، $أ = \overline{ص} - ب\,\overline{س}$