جامعة جيجل

Université de Jijel

كلية العلوم الإنسانية والاجتماعية

قسم علم الاجتماع

مقياس الإحصاء التطبيقي في البحوث الاجتماعية

الإنحدار الخطي البسيط

Simple Linear Regression

إعداد الدكتورة: سامية بوكحيل

السنة الجامعية: 2024 - 2025

مقدمة: ما هو الإنحدار الخطي البسيط؟

يُعد الإنحدار الخطي البسيط (Simple Linear Regression) أحد أهم وأكثر الأساليب الإحصائية استخداماً في البحث العلمي التطبيقي. وهو أسلوب إحصائي يُستخدم لدراسة طبيعة العلاقة بين متغيرين، أحدهما يُسمى المتغير المستقل (Independent Variable) أو متغير التنبؤ، والآخر يُسمى المتغير التابع (Dependent Variable) أو متغير الاستجابة. يهدف الإنحدار الخطي إلى إيجاد معادلة خطية رياضية تُصف العلاقة بين المتغيرين، مما يُمكّن الباحث من التنبؤ بقيم المتغير التابع بناءً على معرفة قيم المتغير المستقل.

يختلف الإنحدار الخطي عن معامل الارتباط في أن الارتباط يقيس قوة العلاقة فقط، بينما الإنحدار يُنتج معادلة رياضية تُمكّن من التنبؤ. فإذا كان معامل ارتباط بيرسون يُجيب على سؤال "هل توجد علاقة بين المتغيرين وكيف قوتها؟"، فإن الإنحدار الخطي يُجيب على سؤال إضافي هو "ما هي المعادلة التي تصف هذه العلاقة وكيف يمكنني التنبؤ بقيم أحد المتغيرين من الآخر؟". هذا التمييز يجعل الإنحدار الخطي أداة أقوى وأكثر فائدة في التطبيقات البحثية العملية.

● **تعريف الإنحدار الخطي البسيط:** هو أسلوب إحصائي يهدف إلى إيجاد أفضل خط مستقيم يُمثّل العلاقة بين متغير مستقل واحد (X) ومتغير تابع واحد (Y)، وذلك من خلال معادلة على شكل: $Y = a + bX$، حيث Y هو المتغير التابع، X هو المتغير المستقل، a هو التقاطع، و b هو ميل الخط.

● **الفرق عن الارتباط:** الارتباط يقيس قوة العلاقة واتجاهها، أما الإنحدار فيُنتج معادلة تُمكّن من التنبؤ بقيم المتغير التابع. الارتباط متماثل ($r(XY) = r(YX)$)، لكن الإنحدار غير متماثل (انحدار Y على X يختلف عن انحدار X على Y).

محتويات المحاضرة

1. أهمية الإنحدار الخطي في البحث العلمي

2. مفهوم المتغير المستقل والمتغير التابع
3. فرضيات الإنحدار الخطي البسيط
4. معادلة خط الإنحدار ومكوناتها
5. طريقة المربعات الصغرى OLS
6. حساب معاملات الإنحدار (a و b)
7. تطبيق عملي مفصّل خطوة بخطوة
8. معامل التحديد R^2 وتفسيره
9. التنبؤ باستخدام معادلة الإنحدار
10. تقييم جودة نموذج الإنحدار
11. الفرق بين انحدار Y على X وانحدار X على Y
12. التطبيق باستخدام برنامج SPSS
13. تطبيقات الإنحدار في العلوم الاجتماعية

أولاً: أهمية الإنحدار الخطي في البحث العلمي

يحتل الإنحدار الخطي مكانة محورية في المنهجية البحثية الكمية، حيث يُعد الأساس الذي تُبنى عليه أساليب إحصائية أكثر تقدماً مثل الإنحدار المتعدد والتحليل التمييزي وتحليل المسار. إن قدرته على التنبؤ ووصف العلاقات الكمّية تجعله أداة لا غنى عنها للباحث في مختلف المجالات العلمية.

1. التنبؤ بالظواهر المستقبلية

يُعد التنبؤ من أهم الوظائف التي يُقدمها الإنحدار الخطي. فمن خلال معادلة الانحدار، يمكن للباحث التنبؤ بقيم المتغير التابع عندما تكون قيمة المتغير المستقل معروفة. على سبيل المثال،

يمكن التنبؤ بتحصيل الطالب بناءً على عدد ساعات الدراسة، أو التنبؤ بمستوى الرضا الوظيفي بناءً على الراتب، أو التنبؤ بمعدل الجريمة بناءً على مستوى البطالة. هذا التنبؤ ليس مطلقاً بل احتمالي، حيث يأخذ بعين الاعتبار هامش الخطأ.

2. قياس الأثر الكمي للمتغيرات

يُمكّن الإنحدار الخطي الباحث من قياس مقدار التغيّر في المتغير التابع عندما يتغيّر المتغير المستقل بمقدار وحدة واحدة. فمعامل الميل (b) في معادلة الانحدار يُعطي هذه القياسات الكمية الدقيقة. على سبيل المثال، إذا كانت معادلة الانحدار للتحصيل على الدراسة هي: $Y = 3 + 1.5X$، فهذا يعني أن كل ساعة دراسة إضافية تزيد التحصيل بمقدار 1.5 درجة. هذا القياس الكمي له أهمية بالغة في صياغة السياسات واتخاذ القرارات المبنية على الأدلة.

3. اختبار النظريات والفرضيات

يُستخدم الإنحدار الخطي في اختبار الفرضيات البحثية بطريقة كمية دقيقة. فمن خلال فحص دلالة معامل الميل (b)، يمكن للباحث قبول أو رفض الفرضية البحثية. كما يُتيح حساب معامل التحديد (R^2) تقييم مدى قدرة المتغير المستقل على تفسير التباين في المتغير التابع، مما يُساعد في الحكم على قوة النظرية المُفترضة.

4. التحكم في المتغيرات والوصف الدقيق

رغم أن الإنحدار البسيط يتعامل مع متغير مستقل واحد فقط، إلا أنه يشكل الأساس لفهم الإنحدار المتعدد الذي يتحكم في عدة متغيرات. كذلك يُوفر الإنحدار الخطي وصفاً دقيقاً للعلاقة من خلال خط الانحدار الذي يُلخّص العلاقة بين المتغيرين في معادلة واحدة واضحة.

ثانياً: مفهوم المتغير المستقل والمتغير التابع

قبل الدخول في تفاصيل الإنحدار الخطي، من الضروري فهم الفرق بين نوعي المتغيرات الرئيسيين في التحليل: المتغير المستقل والمتغير التابع. هذا التمييز هو حجر الأساس في أي تحليل إنحداري، ويؤثر بشكل مباشر على تفسير النتائج.

المتغير المستقل (X) - Independent Variable

هو المتغير الذي يُفترض أنه يؤثر في المتغير التابع أو يُنبّئ به. يُسمى أيضاً المتغير المُفسّر (Explanatory Variable) أو متغير التنبؤ (Predictor Variable) أو السبب المُفترض. في معادلة $Y = a + bX$، يُمثّل X المتغير المستقل. يُعتبَر هذا المتغير هو الذي يتحكم فيه الباحث أو يُلاحظه لدراسة أثره. أمثلة: عدد ساعات الدراسة، مستوى الدخل، العمر، مستوى التعليم.

المتغير التابع (Y) - Dependent Variable

هو المتغير الذي نتوقع أن يتأثر بالمتغير المستقل أو الذي نرغب في التنبؤ بقيمه. يُسمى أيضاً المتغير المُستجيب (Response Variable) أو المتغير المُفسَّر (Explained Variable) أو الأثر المُفترض. في معادلة $Y = a + bX$، يُمثّل Y المتغير التابع. هذا المتغير هو ما يهتم الباحث بفهمه وشرحه. أمثلة: التحصيل الدراسي، مستوى الرضا، الأداء الوظيفي، درجة القلق.

أمثلة تطبيقية

- **في التربية:** المتغير المستقل = ساعات المذاكرة (X) ← المتغير التابع = درجة الامتحان (Y)
- **في علم الاجتماع:** المتغير المستقل = الدخل الشهري (X) ← المتغير التابع = مستوى الرضا عن الحياة (Y)
- **في علم النفس:** المتغير المستقل = عدد ساعات النوم (X) ← المتغير التابع = مستوى التوتر (Y)
- **في الاقتصاد:** المتغير المستقل = مستوى السعر (X) ← المتغير التابع = كمية الطلب (Y)

⚠ **تحذير مهم:** تحديد أي متغير مستقل وأي متغير تابع يعتمد على الإطار النظري للدراسة وليس على الإحصاء. الإنحدار نفسه لا يُميّز بين السبب والنتيجة، بل هذه المهمة يقع على عاتق الباحث بناءً على النظرية والمنطق العلمي.

ثالثاً: فرضيات الإنحدار الخطي البسيط

يعتمد الإنحدار الخطي البسيط على مجموعة من الفرضيات (Assumptions) التي يجب التحقق من تحققها لضمان صحة النتائج وموثوقيتها. انتهاك هذه الفرضيات قد يؤدي إلى تقديرات مُتحيّزة ودلالات إحصائية غير دقيقة.

1. الخطية (Linearity)

الفرضية الأساسية هي أن العلاقة بين المتغير المستقل والمتغير التابع خطية، أي أن التغيّر في Y يكون ثابتاً لكل وحدة تغيّر في X. يمكن فحص هذه الفرضية بصرياً من خلال مخطط الانتشار (Scatter Plot): إذا كانت النقاط تتوزع حول خط مستقيم بشكل عام، فإن الفرضية متحققة. أما إذا كانت تتوزع بشكل منحني، فيجب استخدام نماذج إنحدار غير خطية.

2. الاستقلالية (Independence of Errors)

يجب أن تكون الأخطاء (التباينات) مستقلة عن بعضها البعض، أي أن خطأ المشاهدة i لا يرتبط بخطأ المشاهدة j. انتهاك هذه الفرضية يُسمى الارتباط التسلسلي (Autocorrelation) وهو شائع في البيانات الزمنية. يمكن فحصه باستخدام اختبار دوربين-واتسون (Durbin-Watson test).

3. التجانس في التباين (Homoscedasticity)

يجب أن يكون تباين الأخطاء ثابتاً عند جميع مستويات المتغير المستقل X. أي أن انتشار النقاط حول خط الانعدار يجب أن يكون متساوياً تقريباً عند قيم X المختلفة. إذا كان التباين يتغير (يزداد أو ينقص) مع تغيّر X، يُسمى ذلك عدم تجانس التباين (Heteroscedasticity) ويجب معالجته.

4. التوزيع الطبيعي للأخطاء (Normality of Errors)

يُفترض أن تتبع الأخطاء (الفرق بين القيم الفعلية والقيم المتنبأ بها) توزيعاً طبيعياً بمتوسط صفري. يمكن فحص ذلك باستخدام مخطط Q-Q أو اختبار شابيرو-ويلك. إذا لم يتحقق هذا الشرط، قد تكون الدلالة الإحصائية لمعاملات الإنحدار غير موثوقة عند حجم العينات الصغيرة.

5. عدم وجود تعدّي خطي (No Multicollinearity)

في الإنحدار البسيط (متغير مستقل واحد)، لا تنطبق هذه الفرضية بشكل مباشر، لكنها تصبح حاسمة في الإنحدار المتعدد. مع ذلك يجب ملاحظة أن المتغير المستقل يجب ألا يكون ثابتاً (لا تباين له) لأن ذلك يمنع تقدير معامل الميل.

الفرضية	المعنى	طريقة الفحص	عاقبة الانتهاك
الخطية	العلاقة خطية	مخطط الانتشار	نموذج غير ملائم
الاستقلالية	الأخطاء مستقلة	اختبار Durbin-Watson	تقديرات مُتحيّزة
تجانس التباين	تباين الأخطاء ثابت	مخطط الأخطاء المُتنبأ بها	انحراف معايير الأخطاء
طبيعية الأخطاء	الأخطاء طبيعية التوزيع	مخطط Q-Q / Shapiro-Wilk	دلالة إحصائية غير موثوقة

رابعاً: معادلة خط الإنحدار ومكوناتها

معادلة خط الإنحدار الخطي البسيط هي المعادلة الرياضية التي تصف العلاقة بين المتغير المستقل X والمتغير التابع Y بأفضل طريقة ممكنة. تأخذ هذه المعادلة الشكل العام التالي:

$$Y = a + bX$$

حيث كل مُكوّن له دلالة إحصائية محددة:

المكوّنات الأساسية

1. المتغير التابع (Y):

يُمثّل القيمة المتنبأ بها أو المُقدَّرة للمتغير التابع عند قيمة معينة من المتغير المستقل X. يُسمى أيضاً القيمة المُنسّبة إلى خط الانحدار (Fitted Value) أو القيمة المتوقعة (Expected Value).

ويُرمز لها أحياناً بـ (Y-hat) Ŷ للتمييز بينها وبين القيمة الفعلية Y.

2. التقاطع (a - Intercept):

يُمثّل قيمة Y عندما تكون X = 0 (نقطة تقاطع خط الانحدار مع المحور العمودي). يدل على القيمة المتوقعة للمتغير التابع عندما ينعدم المتغير المستقل تماماً. وقد يكون للتقاطع تفسير منطقي في بعض الحالات (مثلاً: الراتب عندما تكون الخبرة = 0)، وقد لا يكون له تفسير منطقي في حالات أخرى (مثلاً: التحصيل عندما تكون ساعات الدراسة = 0).

3. الميل (b - Slope / Regression Coefficient):

يُمثّل مقدار التغيّر في Y لكل وحدة تغيّر واحدة في X. وهو أهم مكوّن في المعادلة لأنه يُقيس أثر المتغير المستقل في المتغير التابع. إذا كان b موجباً، فهذا يعني أن العلاقة طردية (زيادة X تزيد Y). وإذا كان b سالباً، فالعلاقة عكسية (زيادة X تنقص Y). وإذا كان b = 0، فهذا يعني عدم وجود علاقة خطية.

```
b = ΔY / ΔX (تغيّر Y مقسوماً على مقدار تغيّر X)
```

● **مثال مُبسّط:** إذا كانت معادلة الانحدار: التحصيل = 5 + 2 × (ساعات الدراسة)، فهذا يعني:

- التقاطع (a) = 5: الطالب الذي لا يدرس أبداً (0 ساعات) يُتوقع أن يحصل على 5 درجات.
- الميل (b) = 2: كل ساعة دراسة إضافية تزيد التحصيل بمقدار 2 درجات.
- التنبؤ: طالب يدرس 6 ساعات: $Y = 5 + (2 \times 6) = 17$ درجة.

خامساً: طريقة المربعات الصغرى OLS

طريقة المربعات الصغرى العادية (Ordinary Least Squares - OLS) هي الطريقة الأكثر شيوعاً لتقدير معاملات الإنحدار a و b. تعمل هذه الطريقة على إيجاد الخط الذي يُقلّص مجموع مربعات الفروق بين القيم الفعلية والقيم المُتنبأ بها إلى أدنى حد ممكن.

الفكرة الأساسية

لكل مشاهدة في البيانات، هناك فرق بين القيمة الفعلية الفعلية Y(i) والقيمة المُتنبأ بها Ŷ(i) من خط الانحدار. هذا الفرق يُسمى الخطأ أو المُتبقية (Residual) ويُرمز له بـ e(i):

$$e(i) = Y(i) - \hat{Y}(i) = Y(i) - (a + bX(i))$$

تهدف طريقة المربعات الصغرى إلى إيجاد قيم a و b التي تُقلّص مجموع مربعات هذه الأخطاء:

$$\text{Minimize } \Sigma\, e^2(i) = \Sigma\, [Y(i) - (a + bX(i))]^2$$

لماذا المربعات وليس المطلقات؟

يُستخدم مجموع المربعات بدلاً من مجموع المطلقات لأسباب رياضية متعددة: أولاً، المربعات تُعطي وزناً أكبر للأخطاء الكبيرة مما يُحسّن دقة التقدير. ثانياً، المربعات تجعل الدالة الرياضية قابلة للتفاضل مما يُسهّل إيجاد الحل الأمثل. ثالثاً، المربعات لها خصائص إحصائية مرغوبة كالتقدير غير المُتحيّز والكفاءة.

مخططي التصور

تخيّل مخطط الانتشار الذي يُظهر نقاط البيانات: كل نقطة تُمثّل فرداً في العينة بقيمته في X و Y. خط الإنحدار يمر بين هذه النقاط بحيث يكون مجموع المسافات العمودية (المربعة) بين كل نقطة والخط هو الأصغر الممكن. بعض النقاط تقع فوق الخط (أخطاء موجبة) وبعضها تحته (أخطاء سالبة)، لكن مجموع المربعات يُمثّل إجمالي عدم الدقة.

● **خاصية مهمة:** باستخدام طريقة OLS، يمر خط الانحدار دائماً بنقطة المتوسطات ($\bar{X}$, $\bar{Y}$). هذه الخاصية تضمن أن خط الإنحدار يُمثّل "مرآز" البيانات بشكل صحيح. كما أن مجموع الأخطاء (بدون تربيع) يساوي دائماً صفراً: $\Sigma\, e(i) = 0$.

سادساً: حساب معاملات الإنحدار (a و b)

بعد فهم طريقة المربعات الصغرى، يمكننا الانتقال إلى الصيغ الرياضية الفعلية لحساب معاملي الإنحدار a و b. يوجد أكثر من صيغة مُكافئة، وسنعرض الأكثر استخداماً:

1. حساب الميل (b)

$$b = [n(\Sigma XY) - (\Sigma X)(\Sigma Y)] / [n(\Sigma X^2) - (\Sigma X)^2]$$

حيث n هو حجم العينة، ΣXY هو مجموع حاصل ضرب قيم X في قيم Y لكل مشاهدة، ΣX و ΣY هما مجموعات القيم لكل متغير، و ΣX^2 هو مجموع مربعات قيم X. هذه المعادلة تُمثّل نسبة التغاير بين X و Y إلى تباين X.

الصيغة البديلة باستخدام الانحرافات

$$b = \Sigma[(X(i) - \bar{X})(Y(i) - \bar{Y})] / \Sigma[(X(i) - \bar{X})^2] = S(xy) / S^2(x)$$

حيث S(xy) هو التغاير بين X و Y، و $S^2(x)$ هو تباين X. هذه الصيغة تُبيّن بوضوح أن الميل هو التغاير مقسوماً على التباين.

العلاقة بين الميل ومعامل ارتباط بيرسون

$$b = r(xy) \times (Sy / Sx)$$

هذه العلاقة المهمة تُظهر أن الميل يساوي معامل ارتباط بيرسون مضروباً في نسبة الانحراف المعياري لـ Y إلى الانحراف المعياري لـ X. هذه المعادلة تُبيّن أن الميل يعتمد على وحدات القياس، على عكس معامل الارتباط الذي يخلو من الوحدات.

2. حساب التقاطع (a)

$$a = \bar{Y} - b\bar{X}$$

بمجرد حساب الميل b، يُحسب التقاطع a بسهولة عن طريق طرح حاصل ضرب الميل في متوسط X من متوسط Y. هذه المعادلة تعكس خاصية مرور خط الإنحدار بنقطة المتوسطات.

$$Y = a + bX \longleftrightarrow \bar{Y} = a + b\bar{X} \text{ (عند المتوسطات)}$$

سابعاً: تطبيق عملي مفصّل خطوة بخطوة

للتوضيح العملي لطريقة حساب معادلة الإنحدار الخطي البسيط، سنستخدم مثالاً يتناول العلاقة بين عدد ساعات الدراسة الأسبوعية (X) ودرجة التحصيل من 20 (Y) لدى عينة من 5 طلاب.

البيانات

الطالب	ساعات الدراسة X	التحصيل Y	X^2	Y^2	XY
1	3	7	9	49	21
2	5	11	25	121	55
3	2	5	4	25	10
4	8	15	64	225	120
5	6	12	36	144	72
المجموع	**$\Sigma X = 24$**	**$\Sigma Y = 50$**	**$\Sigma X^2 = 138$**	**$\Sigma Y^2 = 564$**	**$\Sigma XY = 278$**

خطوات الحل

1 **حساب المتوسطات:** $\bar{X} = 24/5 = 4.8$ ساعة، $\bar{Y} = 50/5 = 10$ درجات

2 **حساب الميل (b):**

$$b = [n(\Sigma XY) - (\Sigma X)(\Sigma Y)] / [n(\Sigma X^2) - (\Sigma X)^2]$$

$$b = [(5 \times 278) - (24 \times 50)] / [(5 \times 138) - (24)^2]$$

$$b = [1390 - 1200] / [690 - 576] = 190 / 114 = 1.667$$

3 **حساب التقاطع (a):**

$$a = \bar{Y} - b\bar{X} = 10 - (1.667 \times 4.8) = 10 - 8 = 2$$

4 **معادلة خط الإنحدار:**

$$\hat{Y} = 2 + 1.667X$$

✓ التفسير:

• **التقاطع (a = 2):** عندما تكون ساعات الدراسة صفراً، يُتوقع أن يكون التحصيل 2 درجات (وهو أمر افتراضي).

• **الميل ($b = 1.667$):** كل ساعة دراسة إضافية تزيد التحصيل بمقدار 1.667 درجة تقريباً.

• **التنبؤ:** طالب يدرس 7 ساعات أسبوعياً: $\hat{Y} = 2 + (1.667 \times 7) = 13.67$ درجة.

• **التنبؤ:** طالب يدرس 4 ساعات: $\hat{Y} = 2 + (1.667 \times 4) = 8.67$ درجة.

جدول التقييم (القيم المتنبأ بها والأخطاء)

الطالب	X	Y الفعلي	$\hat{Y}$ المتنبأ	$e = Y - \hat{Y}$	e^2
1	3	7	7.00	0.00	0.00
2	5	11	10.33	0.67	0.44
3	2	5	5.33	0.33-	0.11
4	8	15	15.33	0.33-	0.11
5	6	12	12.00	0.00	0.00
المجموع				**0 ≈**	**0.66**

ثامناً: معامل التحديد R^2 وتفسيره

معامل التحديد (Coefficient of Determination) R^2 هو مقياس إحصائي يُبيّن مدى قدرة خط الانحدار على تفسير التباين في المتغير التابع Y. يتراوح بين 0 و 1، ويُفسَّر كنسبة مئوية. يُعد هذا المعامل من أهم المؤشرات لتقييم جودة نموذج الإنحدار.

المعادلة

```
R² = SSR / SST = 1 - (SSE / SST)
```

حيث:

- **SST (Total Sum of Squares):** مجموع المربعات الكلي = $\Sigma(Y(i) - \bar{Y})^2$ ← إجمالي التباين في Y
- **SSR (Regression Sum of Squares):** مجموع مربعات الانحدار = $\Sigma(\hat{Y}(i) - \bar{Y})^2$ ← التباين المُفسَّر
- **SSE (Error Sum of Squares):** مجموع مربعات الأخطاء = $\Sigma(Y(i) - \hat{Y}(i))^2$ ← التباين غير المُفسَّر

```
SST = SSR + SSE (إجمالي التباين = التباين المُفسّر + التباين غير المُفسّر)
```

التفسير

قيمة R^2	المعنى	التقييم
$R^2 = 0.00$	خط الانحدار لا يُفسّر أي تباين في Y	نموذج فاشل تماماً
$R^2 = 0.25$	25% من التباين مُفسَّر و75% غير مُفسَّر	ضعيف
$R^2 = 0.50$	50% من التباين مُفسَّر	متوسط
$R^2 = 0.75$	75% من التباين مُفسَّر	جيد
$R^2 = 1.00$	100% من التباين مُفسَّر (نقاط على الخط)	ملاءمة مثالية

العلاقة بين R^2 ومعامل الارتباط

في الإنحدار الخطي البسيط (متغير مستقل واحد)، يكون معامل التحديد مساوياً لمربع معامل ارتباط بيرسون:

```
R² = r²(xy) (في الإنحدار البسيط فقط)
```

● **في مثالنا السابق:** $r = 0.996$ (ارتباط شبه مثالي)، وبالتالي $R^2 = (0.996)^2 = 0.992$. هذا يعني أن ساعات الدراسة تُفسّر 99.2% من التباين في التحصيل الدراسي، وهي نسبة مرتفعة جداً تدل على ملاءمة النموذج الممتازة.

⚠ **تنبيه:** في العلوم الاجتماعية، قيمة $R^2 = 0.30$ تعتبر جيدة عادة. فالسلوك الإنساني يخضع لتأثير متغيرات كثيرة، والحصول على R^2 مرتفع نادر. لا تُقيّم R^2 بمعايير مطلقة بل في سياق المجال البحثي وطبيعة المتغيرات المدروسة.

تاسعاً: التنبؤ باستخدام معادلة الإنحدار

من أهم الاستخدامات العملية لمعادلة الإنحدار هي التنبؤ بقيم المتغير التابع عندما تكون قيمة المتغير المستقل معروفة. يُعد هذا التطبيق مفيداً في كثير من المجالات كالتربوية والاجتماعية والاقتصادية.

أنواع التنبؤ

1. التنبؤ الداخلي (Interpolation)

هو التنبؤ بقيم Y عند قيم X تقع ضمن نطاق البيانات الأصلية. هذا النوع من التنبؤ أكثر موثوقية لأنه يعتمد على بيانات فعلية مُلاحظة. على سبيل المثال، إذا كانت بيانات ساعات الدراسة في العينة تتراوح بين 2 و 8 ساعات، فإن التنبؤ لطالب يدرس 5 ساعات يُعتبر تنبؤاً داخلياً موثوقاً.

2. التنبؤ الخارجي (Extrapolation)

هو التنبؤ بقيم Y عند قيم X تقع خارج نطاق البيانات الأصلية. هذا النوع محفوف بالمخاطر لأنه يفترض استمرار نفس العلاقة الخطية خارج النطاق المُلاحظ، وهو افتراض قد لا يكون صحيحاً. على سبيل المثال، التنبؤ بتحصيل طالب يدرس 20 ساعة (بينما أقصى قيمة في العينة 8) قد يُعطي نتائج غير منطقية.

التنبؤ الاحتمالي وهامش الخطأ

التنبؤ بمعادلة الانحدار لا يُعطي قيمة مؤكدة، بل قيمة متوقعة مع هامش خطأ. يتم التعبير عن ذلك بفترات الثقة (Confidence Intervals):

$$\hat{Y} \pm t(\alpha/2, n-2) \times Se \times \sqrt{[1 + 1/n + (X_0 - \bar{X})^2 / \Sigma(X(i) - \bar{X})^2]}$$

حيث Se هو الخطأ المعياري لل estimate، و X_0 هي قيمة X التي نريد التنبؤ عندها، و t هو القيمة الجدولية لتوزيع t عند مستوى الثقة المطلوب. يلاحظ أن هامش الخطأ يزداد كلما ابتعدت قيمة X_0 عن المتوسط $\bar{X}$، وهو أمر منطقي لأن التنبؤ يكون أقل دقة عند حواف البيانات.

مثال تطبيقي

معادلة الانحدار: $\hat{Y} = 2 + 1.667X$

• طالب يدرس 3 ساعات: $\hat{Y} = 2 + (1.667 \times 3) = 7$ درجات

• طالب يدرس 10 ساعات: $\hat{Y} = 2 + (1.667 \times 10) = 18.67$ درجة

• طالب يدرس 0 ساعات: $\hat{Y} = 2 + 0 = 2$ درجة

ملاحظة: التنبؤ عند X = 10 ساعات (خارج نطاق العينة 2-8) يجب التعامل معه بحذر.

عاشراً: تقييم جودة نموذج الإنحدار

بعد تقدير معادلة الإنحدار، يجب تقييم جودة النموذج من عدة جوانب لتحديد مدى ملاءمته وموثوقية نتائجه. يتضمن التقييم عدة مؤشرات إحصائية واختبارات ضرورية.

1. الخطأ المعياري للتقدير (Standard Error of Estimate - Se)

$$Se = \sqrt{[SSE / (n - 2)]} = \sqrt{[\Sigma(Y - \hat{Y})^2 / (n - 2)]}$$

يُقيس متوسط الانحرافات للقيم الفعلية عن القيم المُتنبأ بها. كلما كان Se أصغر، كانت دقة التنبؤ أعلى. إذا كان Se صغيراً نسبياً مقارنة بمتوسط Y، فهذا يدل على نموذج جيد. يمكن تفسير Se بأنه متوسط الخطأ الذي نتوقعه عند التنبؤ بقيم Y باستخدام النموذج.

2. اختبار دلالة الميل (t-test for slope)

$$t = b / Sb \text{ where } Sb = Se / \sqrt{[\Sigma(X - \bar{X})^2]}$$

يختبر هذا الاختبار ما إذا كان الميل b دالاً إحصائياً (مختلفاً عن الصفر). الفرضية العدمية هي: H_0: B = 0 (لا توجد علاقة خطية). إذا كانت قيمة t المحسوبة أكبر من القيمة الجدولية (أو كانت قيمة Sig. أقل من 0.05)، نرفض الفرضية العدمية ونستنتج أن الميل دال إحصائياً.

3. اختبار دلالة النموذج (F-test)

$$F = MSR / MSE = [SSR/1] / [SSE/(n-2)]$$

يختبر ما إذا كان النموذج ككل دالاً إحصائياً (أي هل المتغير المستقل يُفسّر جزءاً معنوياً من تباين Y). في الإنحدار البسيط، اختبار F يُعادل اختبار t للميل، لكنه يصبح مختلفاً وأكثر أهمية في الإنحدار المتعدد.

4. فحص المُتبقيات (Residual Analysis)

يُعد فحص المُتبقيات (الفروق بين القيم الفعلية والمُتنبأ بها) وسيلة مهمة للتأكد من تحقق فرضيات الإنحدار. يتم ذلك من خلال:

- **مخطط المُتبقيات مقابل القيم المُتنبأ بها:** يجب أن يكون التوزيع عشوائياً بدون نمط واضح (للتحقق من تجانس التباين).

- **مخطط Q-Q للمُتبقيات:** يجب أن تقع النقاط على الخط القطري تقريباً (للتحقق من التوزيع الطبيعي).
- **مخطط المُتبقيات مقابل الترتيب:** يجب أن يكون التوزيع عشوائياً (للتحقق من الاستقلالية).

حادي عشر: الفرق بين انحدار Y على X وانحدار X على Y

من المفاهيم المهمة في الإنحدار الخطي أن انحدار Y على X يختلف عن انحدار X على Y، أي أن معادلتين مختلفتين تنتجان عن نفس البيانات حسب اختيار المتغير التابع والمتغير المستقل. هذا الفرق أساسي ويجب فهمه بدقة.

انحدار Y على X

$$\hat{Y} = a(yx) + b(yx) \times X$$

$$b(yx) = r \times (Sy / Sx)$$

يُستخدم عندما نريد التنبؤ بقيم Y من معرفة X. يُقلّص هذا الخط مجموع مربعات الفروق العمودية (أي الفروق في اتجاه محور Y).

انحدار X على Y

$$\hat{X} = a(xy) + b(xy) \times Y$$

$$b(xy) = r \times (Sx / Sy)$$

يُستخدم عندما نريد التنبؤ بقيم X من معرفة Y. يُقلّص هذا الخط مجموع مربعات الفروق الأفقية (أي الفروق في اتجاه محور X).

جدول المقارنة

وجه المقارنة	انحدار Y على X	انحدار X على Y
الهدف	التنبؤ بـ Y من X	التنبؤ بـ X من Y
المتغير التابع	Y	X
المتغير المستقل	X	Y
الميل	$b(yx) = r(Sy/Sx)$	$b(xy) = r(Sx/Sy)$
يُقلّص	الفروق العمودية ($\Sigma e^2 y$)	الفروق الأفقية ($\Sigma e^2 x$)
معادلة الميل	$b(yx) \times b(xy) = r^2$	نفس العلاقة
التقاطع	$a = \bar{Y} - b\bar{X}$	$a = \bar{X} - b\bar{Y}$

● **متى تتقاطع الخطين؟** خطوط الانحدار (Y على X و X على Y) تتقاطع دائماً عند نقطة المتوسطات ($\bar{X}, \bar{Y}$). كلما كان معامل الارتباط أقوى (أقرب إلى 1 أو -1)، اقترب الخطين من بعضهما. وعندما $r = \pm 1$، يتطابق الخطين تماماً.

ثاني عشر: التطبيق باستخدام برنامج SPSS

يُوفر برنامج SPSS إمكانية حساب معادلة الإنحدار الخطي البسيط بسهولة من خلال خطوات محددة، مع تقديم مخرجات تفصيلية تشمل معاملات الإنحدار ومعامل التحديد واختبارات الدلالة وتحليل المُتبقيات.

خطوات الحساب في SPSS

1 **إدخال البيانات:** نُدخل بيانات المتغير المستقل (X) والمتغير التابع (Y) في عمودين منفصلين في محرر البيانات. نحدد نوع كل متغير كـ "Scale" في نافذة Variable View.

2 **فتح نافذة التحليل:** من القائمة: Analyze → Regression → Linear

3 **تحديد المتغيرات:** نُنقل المتغير التابع (Y) إلى مربع Dependent، والمتغير المستقل (X) إلى مربع Independent(s). Method نتركها "Enter".

4 **الخيارات الإحصائية:** من زر Statistics نختار: Estimates (معاملات الإنحدار)، Model fit (R^2 واختبار F)، Descriptives (المتوسطات والانحرافات)، و Collinearity diagnostics.

5 **مخططات المُتبقيات:** من زر Plots نختار: Histogram (لتوزيع المُتبقيات)، Normal probability plot (للتحقق من الطبيعية)، و ZRESID* مقابل ZPRED* (للتحقق من تجانس التباين).

6 **تشغيل التحليل:** نضغط OK للحصول على المخرجات.

قراءة أهم المخرجات

الجدول	المؤشر	المعنى
Model Summary	R, R^2, Adj R^2, Std. Error	قوة النموذج وملاءمته
ANOVA	F, Sig.	دلالة النموذج ككل
Coefficients	B (المعامل), t, Sig.	دلالة كل معامل إنحدار

Residuals Statistics	Min, Max, Mean, Std.Dev.	ملخص المُتبقيات

✓ **مثال على تفسير مخرجات SPSS:**

• **Model Summary:** $R = 0.98$, $R^2 = 0.96$, Std. Error = 0.72

← ارتباط قوي جداً، النموذج يُفسّر 96% من التباين.

• **ANOVA:** $F = 245.3$, Sig. = 0.000

← النموذج دال إحصائياً عند مستوى 0.001.

• **Coefficients:** (Constant) $B = 2.00$, Sig. = 0.003 | (X) $B = 1.67$, Sig. = 0.000

← التقاطع = 2 والميل = 1.67، كلاهما دال إحصائياً.

• **المعادلة النهائية:** $\hat{Y} = 2.00 + 1.67X$

ثالث عشر: تطبيقات الإنحدار في العلوم الاجتماعية

يُستخدم الإنحدار الخطي البسيط على نطاق واسع في مختلف فروع العلوم الاجتماعية، حيث يُساهم في فهم العلاقات الكمية بين المتغيرات والتنبؤ بالظواهر الاجتماعية. فيما يلي أبرز التطبيقات في كل مجال:

1. في علم الاجتماع

- **الحراك الاجتماعي:** دراسة العلاقة بين المستوى التعليمي والدخل، أو بين عدد سنوات الخبرة والمكانة الاجتماعية، مما يُساعد في فهم آليات الحراك الاجتماعي الصعودي والهبوطي.
- **الجريمة والانحراف:** تحليل العلاقة بين معدلات البطالة ومعدلات الجريمة في مناطق مختلفة، أو بين كثافة السكان والحوادث، لصياغة سياسات وقائية أكثر فعالية.

- **المشاركة السياسية:** دراسة العلاقة بين المستوى التعليمي ومستوى المشاركة في الانتخابات، أو بين العمر واتجاهات التصويت.

2. في علم النفس

- **الذكاء والتحصيل:** التنبؤ بالتحصيل الأكاديمي من درجات اختبارات الذكاء، مما يُساعد في التوجيه المدرسي والمهني.
- **الصحة النفسية:** دراسة العلاقة بين مستوى الضغط النفسي والأعراض الاكتئابية، أو بين الدعم الاجتماعي ومستوى التكيّف النفسي.
- **الشخصية:** فحص العلاقة بين أبعاد الشخصية والسلوكيات المختلفة كالأداء الوظيفي أو العلاقات الاجتماعية.

3. في التربية

- **التعلم:** التنبؤ بتحصيل الطلاب بناءً على عدد ساعات الدراسة أو الدافعية أو التحصيل السابق.
- **تقييم البرامج:** قياس أثر البرامج التربوية والتدريبية من خلال مقارنة الأداء قبل وبعد التدخل.
- **صدق الاختبارات:** تقدير صدق المحك للاختبارات من خلال انحدار درجات الاختبار على معيار خارجي.

4. في الاقتصاد

- **الاستهلاك:** دراسة العلاقة بين الدخل ومستوى الاستهلاك (دالة كينز الاستهلاكية).
- **العرض والطلب:** تحليل العلاقة بين السعر والكمية المطلوبة أو المُعرّضة.
- **النمو الاقتصادي:** دراسة العلاقة بين الاستثمار ومعدلات النمو أو بين التعليم والناتج المحلي.

الخلاصة: يُعد الإنحدار الخطي البسيط من أساسيات التحليل الإحصائي التي لا غنى عنها لأي باحث في العلوم الاجتماعية. إن فهم المفاهيم النظرية للإنحدار، من معادلة خطية وطريقة المربعات الصغرى ومعاملات التحديد والتنبؤ، يُمكّن الباحث من وصف العلاقات

بين المتغيرات بدقة كمية والتنبؤ بقيم الظواهر المستقبلية. كما يُشكّل الإنحدار البسيط البوابة نحو الإنحدار المتعدد والتحليل المتقدم الذي يتعامل مع عدة متغيرات مستقلة في آن واحد. يجب على الباحث دائماً التحقق من فرضيات الإنحدار وتفسير النتائج بحذر في ضوء الإطار النظري والسياق البحثي.

المراجع والمصادر

1. Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences (3rd ed.). Lawrence Erlbaum Associates.
2. Field, A. (2018). Discovering Statistics Using IBM SPSS Statistics (5th ed.). SAGE Publications.
3. Gravetter, F. J., & Wallnau, L. B. (2017). Statistics for the Behavioral Sciences (10th ed.). Cengage Learning.
4. Montgomery, D. C., Peck, E. A., & Vining, G. G. (2012). Introduction to Linear Regression Analysis (5th ed.). Wiley.
5. Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Applied Linear Statistical Models (5th ed.). McGraw-Hill.
6. Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). Multivariate Data Analysis (8th ed.). Cengage Learning.
7. Pallant, J. (2020). SPSS Survival Manual (7th ed.). Routledge.
8. أمل سعيد، عبد الرحمن عدس (2010). مقدمة في علم الإحصاء الوصفي والاستدلالي. دار المناهج، عمّان.
9. محمد نجيب الصبّاح (2015). الإحصاء التطبيقي باستخدام SPSS. دار وائل للنشر والتوزيع، عمّان.
10. فيديو: الإنحدار الخطي البسيط - قناة إحصاء على YouTube (قائمة التحليل الإحصائي).