

حل نموذجي لامتحان النهائي في مادة: أدوات الإحصائية وتحليل البيانات

السؤال الأول: تحليل المكونات الرئيسية (Analyse en composantes principales) هي تقنية إحصائية تستخدم في عدة مجالات من تحليل البيانات. هناك العديد من الحالات الشائعة التي يتم فيها تطبيق هذه التقنية نذكر منها على سبيل المثال:
- حالة تقليص البعد (Réduction de Dimension)
- حالة إدارة تعدد الارتباطات (Gestion de la Multicollinéarité)
أ) أشرح هاتين الحالتين (أشرح المفهومين المذكورين)؟

الإجابة

- حالة تقليص البعد
نكون بحاجة ملحة لهذه الحالة عندما يكون لدينا عدد كبير من المتغيرات ونريد تقليل حجم البيانات مع الحفاظ على المعلومات بأفضل شكل ممكن. الهدف هنا يكمن في التقاط أقصى قدر من التباين مع أقل عدد ممكن من المكونات.
- حالة إدارة تعدد الارتباطات
تعد العلاقة الخطية المتعددة ظاهرة تحدث عندما يرتبط متغيران أو أكثر في تحليل المكون الرئيسي ارتباطًا وثيقًا ببعضهما البعض. وهذا يمكن أن يسبب مشاكل لأن هذا التحليل يعتمد على فكرة التقاط المعلومات الموجودة في المتغيرات دون تكرار. عندما تكون هناك علاقة خطية متعددة، قد يكون من الصعب تحديد الأهمية النسبية لكل متغير ويمكن أن يؤثر ذلك على نتائج نوع التحليل. يوصى عمومًا بفحص الارتباط بين المتغيرات قبل إجراء تحليل المكونات الرئيسية (وهذا ما شاهدناه في إحدى الحصص التطبيقية) من أجل اكتشاف العلاقة الخطية المتعددة واتخاذ الخطوات المناسبة للتعامل معها بشكل جيد.
ب) ماذا يحدث لهذه التقنية إذا كانت جل المتغيرات خطية أو لها ارتباطات ضعيفة؟

الإجابة

- إذا كانت جل المتغيرات خطية هذا يؤكد جليا ان تقنية "تحليل المكونات الأساسية" توفر الكثير من المعلومات للالتقاط كل العلاقات الموجودة بين هذه المتغيرات.
- إذا كان للمتغير تباين منخفض جدًا، فلن يكون هناك توفير كافي من المعلومات إلى هذه التقنية فبالتالي لا تعود مجدية.

السؤال الثاني: لدينا تحليلا للمكونات الرئيسية مع ثلاثة مكونات رئيسية (CP1 - CP2 - CP3) وامامنا هذه الجداول الثلاثة:الصادرة من مخرجات المعالجة على أي شيء تدل هذه الجداول وقم بتفسيرها؟

الاجابة

✓ تعريف ماهية الجدول الاول

Composante Principale | Variance Totale Expliquée (%)

CP1		60
CP2		30
CP3		10

في تحليل المكونات الرئيسية، يتم استخدام Variance Totale Expliquée أي "التباين الإجمالي المفسر" لقياس مقدار التباين الإجمالي في البيانات الأصلية الذي يمكن أن يتم تفسيره أو شرحه عن طريق المكونات الرئيسية. يساعدنا ذلك في فهم كم نسبة من التباين في البيانات يمكن أن تشرحه المكونات الرئيسية المختلفة. عند إجراء تحليل المكونات الرئيسية، يتم ترتيب المكونات الرئيسية بحيث يتم وضع أول مكون رئيسي لديه أكبر قدر من "التباين الإجمالي المفسر"، ويتبعه المكون الرئيسي الثاني وهكذا. يساهم كل مكون رئيسي في شرح جزء من التباين الإجمالي في البيانات، والترتيب يعكس أهمية المكونات الرئيسية في شرح التباين (يتم هذا الترتيب عن طريق القيم الذاتية كما جاء في التطبيق). بشكل عام، يحاول تحليل المكونات الرئيسية تقليل الأبعاد الأصلية لمجموعة البيانات وتمثيلها بأبعاد أقل (المكونات الرئيسية). ونسبة التباين المفسر لكل مكونة تساعد في تحديد الأبعاد الأكثر إفادة وتفسيرًا للتباين الكلي للبيانات.

✓ تفسير مضمون الجدول الأول

تشير هذه النسب إلى نسبة التباين الإجمالي الموضح لكل مكون رئيسي. في هذا المثال، يوضح 60% أن المكون الرئيسي CP1 يقوم بشرح وتفسير نسبة 60% من إجمالي التباين نفس الشيء يقال النسبة للمكون CP2 ب 30% وأخيرا CP3 ب 10%.

✓ تعريف ماهية الجدول الثاني

Matrice des Composantes

Variable	CP1	CP2	CP3
Variable_1	0.75	-0.5	0.25
Variable_2	-0.6	-0.8	0.0
Variable_3	0.25	0.3	0.9

تدل مصفوفة المكونات على قوة واتجاه العلاقة بين المتغيرات الأصلية والمكونات الرئيسية. عندما تكون قيم الحمولة كبيرة بشكل ملحوظ لمتغير معين في مكونة محددة، فإن ذلك يشير إلى أن هذا المتغير له تأثير قوي على هذه المكونة. بالإضافة إلى ذلك، يمكن استخدام مصفوفة المكونات لتحديد الأنماط أو العناصر المشتركة بين المتغيرات الأصلية. على سبيل المثال، إذا كانت هناك متغيرات مختلفة مرتبطة بصفات مشتركة في مكونة معينة، فإن ذلك يمكن أن يشير إلى وجود عناصر مشتركة في البيانات تتفق مع تلك الصفات.

بشكل عام، مصفوفة المكونات تعطينا نظرة أكثر تفصيلاً عن العلاقات والتأثيرات بين المتغيرات الأصلية والمكونات الرئيسية في تحليل المكونات الرئيسية. إنها أداة قيمة لفهم البيانات واستخلاص المعلومات المهمة. أتمنى أن يكون ذلك واضحاً، وإذا كانت لديك أي أسئلة أخرى، فأنا في الخدمة.

✓ تفسير مضمون الجدول الثاني

تمثل القيم الموجودة في هذا الجدول الأوزان المخصصة لكل متغير مقابل كل مكون رئيسي. تشير هذه الأوزان إلى مساهمة كل متغير اتجاه كل مكون.

✓ تعريف ماهية الجدول الثالث

Rotation de la Matrice des Composantes

Variable	CP1	CP2	CP3
Variable_1	0.85	0.1	0.5
Variable_2	0.2	-0.9	0.3
Variable_3	0.5	0.4	-0.8

عندما نقوم بتطبيق الدوران على مصفوفة المكونات الرئيسية، فإننا نقوم بتحويل المصفوفة لتكون أكثر تناغمًا وبساطة. تقوم عملية الدوران بتغيير زوايا المحاور في المساحة المكونة، وبالتالي تساهم في تبسيط التفسير وتحسين قابلية الفهم للنتائج. بعد تطبيق الدوران، يمكننا استخدام مصفوفة المكونات المركبات لتحديد العلاقات بين المتغيرات الأصلية والمكونات الرئيسية بشكل أكثر دقة. يمكننا تحليل قيم الحمولة في المصفوفة لتحديد المتغيرات التي تسهم بشكل كبير في كل مكونة بعد الدوران. هذا يمكن أن يساعد في تحديد العوامل المؤثرة الرئيسية وفهم التغيرات المهمة التي تحدث في البيانات.

✓ تفسير مضمون الجدول الثالث

بعد الدوران يمكن أن تتغير الأوزان لتحسين التفسير. تشير الأوزان العالية هنا إلى مساهمة قوية للمتغير في المكون المعني. والشيء الملاحظ من خلال هذا المثال أن القيم الفعلية ستعتمد على

البيانات المرصدة وسيعتمد التفسير الدقيق للنتائج على سياق الدراسة المأخوذة وخصائص المتغيرات.

السؤال الثالث: نفترض أنك تعمل في شركة لإدارة الثروات وتريد تحليل العوامل التي تؤثر على عوائد الاستثمار. ولدراسة هذه العوامل أمامك المتغيرات التالية: العائد على الاستثمار - مبلغ الاستثمار الأولي - سن العميل - مستوى تعليم العميل - عدد سنوات الخبرة في ميدان الاستثمار - درجة المخاطر المرتبطة بالعميل. نفترض من جهة أخرى أن نتائج المعاملات المحققة والتابعة للمتغيرات جاءت كما يلي:

معامل المبلغ الأولي للاستثمار: 0.15 | معامل سن العميل: -0.01
معامل مستوى التعليم: 0.5 | معامل عدد سنوات الخبرة في الاستثمار: 0.02
معامل درجة مخاطر العميل: -0.1 | معامل الاعتراض: 2.5 (ثابت)

(ا) ما هي المسألة المطروحة هنا (ما هي القيمة التي نبحث عنها)؟

الإجابة

نريد في هذه المسألة معرفة فيما إذا كانت المتغيرات المستقلة المعطاة وعددها 5 (مبلغ الاستثمار الأولي - سن العميل - مستوى تعليم العميل - عدد سنوات الخبرة في ميدان الاستثمار - درجة المخاطر المرتبطة بالعميل) لها تأثير على المتغير التابع عوائد الاستثمار أو بمعنى آخر إيجاد قياس المتغير التابع (عوائد الاستثمار) بواسطة المتغيرات المستقلة التي تم ذكرها. تسمح لك هذه التفسيرات بفهم كيفية تأثير كل متغير مستقل على عائد الاستثمار. ومع ذلك، تجدر الإشارة إلى أن هذه التفسيرات تستند إلى البيانات والعينة المحددة المستخدمة في هذا التحليل

(ب) ما هو اسم النموذج المعمول به في هذه المسألة؟

الإجابة

يسمى النموذج المعمول به في هذه المسألة بالانحدار المتعدد

الإجابة

(ج) استنتج معادلة هذا النموذج وما دورها في هذه النمذجة؟

عوائد الاستثمار = 0.15 * العائد على الاستثمار - 0.02 * مبلغ الاستثمار الأولي - 0.01 * سن العميل + 0.5 * مستوى تعليم العميل + 0.02 * عدد سنوات الخبرة في ميدان الاستثمار - 0.1 * درجة المخاطر المرتبطة بالعميل + 2.5

(د) كيف نفسر هذه المعادلة؟

الإجابة

-مبلغ الاستثمار الأولي: لكل زيادة قدرها 1 دينار في مبلغ الاستثمار الأولي، يزداد عائد الاستثمار بمتوسط 0.15%.

- عمر العميل: لكل زيادة لمدة عام في عمر العميل، ينخفض عائد الاستثمار بمعدل 0.01%.
- مستوى تعليم العميل: يتمتع العملاء ذوو المستوى التعليمي العالي، في المتوسط، بعائد استثمار أعلى بنسبة 0.5% مقارنة بالعملاء ذوي المستوى التعليمي المتوسط أو المنخفض.

- عدد سنوات الخبرة في ميدان الاستثمار: مقابل كل زيادة قدرها سنة واحدة من الخبرة الاستثمارية، يزيد عائد الاستثمار بمعدل 0.02%.
- درجة مخاطر العميل: مقابل كل زيادة بمقدار نقطة واحدة في درجة مخاطر العميل، ينخفض عائد الاستثمار بمعدل 0.1%.
- معامل الاعتراض: عندما تكون جميع المتغيرات المستقلة كلها معدومة (تساوي الصفر)، يبلغ متوسط عائد الاستثمار 2.5%.

الاجابة

هـ) ما دور معامل الاعتراض هنا؟

- دور معامل الاعتراض هو قياس قوة واتجاه تأثير المتغير التفسيري على المتغير الاعتمادي بعد التحكم في تأثير المتغيرات الأخرى في النموذج. عندما تكون لدينا معادلة تحليل الانحدار التعددي، يتم تقدير قيمة معامل الاعتراض لكل متغير تفسيري. إذا كان معامل الاعتراض إيجابياً، يشير ذلك إلى أن زيادة في المتغير التفسيري ترتبط بزيادة في المتغير الاعتمادي. وإذا كان معامل الاعتراض سالباً، يشير ذلك إلى أن زيادة في المتغير التفسيري ترتبط بانخفاض في المتغير الاعتمادي. بالإضافة إلى ذلك، يمكن استخدام معامل الاعتراض لتحديد أهمية المتغيرات التفسيرية المختلفة في توقع المتغير الاعتمادي. فعندما يكون لديك معامل اعتراض كبير لمتغير تفسيري معين، يعني ذلك أن هذا المتغير يلعب دوراً مهماً في توقع المتغير الاعتمادي.
- بشكل عام، يساعد معامل الاعتراض في تحليل الانحدار التعددي على فهم العلاقة بين المتغيرات وتوقع القيم المستقبلية للمتغير الاعتمادي. كما يساعد في تحديد التأثيرات الفردية للمتغيرات التفسيرية وتوفير رؤية شاملة للعلاقات بين المتغيرات في النموذج.
- و) نحاول ان نعالج هذه المسألة عن طريق SPSS أعط المسار المؤدي للوصول الى النموذج المشار إليه في السؤال (ا-1)؟

الاجابة

- افتح SPSS وقم بتحميل ملف البيانات الموجه للتحليل
- اتبع هذا المسار Analyze/ Regression/ Linear
- قم بادخل المتغير التابع في المربع dependent
- والمتغيرات المستقلة في المربع independents
- ملاحظة: بالنسبة لنسخة SPSS بالفرنسي نأخذ ما يلي
- Analyse/ Régression/ Linéaire
- قم بادخل المتغير التابع في المربع dépendant
- والمتغيرات المستقلة في المربع indépendants

السؤال الرابع: نقترح عليك 4 تطبيقات، ما اسم نموذج الاختبار الاحصائي المناسب للاستعمال في كل تطبيق وكيف نصل اليه في نظام SPSS؟ (لا نطلب منك إعطاء الحل).

الاجابة

التطبيق الاول: توقع اتجاهات سوق الأسهم

يمكن استخدام نموذج الانحدار المتعدد. هذا النموذج يمكنه تحليل العلاقة بين متغير استجابة (توقعات سوق الأسهم) ومتغيرات تفسيرية (مؤشرات السوق وعوامل أخرى ذات صلة).

التطبيق الثاني: نمذجة المخاطر المالية واتخاذ قرارات صائبة بشأن إدارة المخاطر وتخصيص الأصول.

يمكن استخدام الانحدار المتعدد. هذا النموذج يمكنه التعامل مع عدة متغيرات تفسيرية في نفس الوقت وتحليل تأثيرها على المخاطر المالية واتخاذ القرارات المناسبة.

بالنسبة للتطبيقين الاول والثاني للوصول الى نموذجهما تتبع هذا المسار

Analyse/ Régression/ Linéaire

التطبيق الثالث: تقييم مخاطر الائتمان للأعمال التجارية عن طريق بناء نموذج تنبؤي لمخاطر الائتمان وتقدير احتمال التخلف عن السداد

يمكن استخدام الانحدار اللوجستي. هذا النموذج يمكنه تصنيف المخاطر الائتمانية وتحليل العوامل المؤثرة في احتمالية التخلف عن السداد.

لوصول لهذا النموذج تتبع هذا المسار **Analyse / Régression / Logistique binaire**
التطبيق الرابع: الرقابة الآلية في تحسين سلامة البيانات المدخلة وجودة المعلومات المحاسبية لضمان الحصول على معلومات محاسبية دقيقة وموثوقة.

يمكن استخدام اختبار كي تربيع. هذا الاختبار يمكنه كشف القيم المتطرفة أو البيانات غير السليمة في مجموعة البيانات ومعالجتها.

لوصول لهذا النموذج تتبع هذا المسار **Statistiques Descriptives / Tableaux croisés**
نختار في صندوق الحوار المعروض **Khi-deux**

مسؤول المادة: علي ابن عافيه

:

