

إن المقصود بمتعدد المتغيرات هو التعامل مع حالة وجود أكثر من متغير واحد سواءً فيما يتعلق بالمتغيرات (العوامل) التوضيحية (وتسمى بالمستقلة أحياناً Independent) من جهة، أو متغيرات الإستجابة (وتسمى بالمعتمدة Dependent) من جهة أخرى.

والبيانات متعددة المتغيرات تظهر في جميع تفرعات العلوم تقريباً. وفي الغالب نجد أن جميع البيانات التي يتم جمعها من خلال الوحدة التجريبية Experimental Unit وتحليلها من قبل الباحثين يمكن تصنيفها بكونها بيانات متعددة المتغيرات. والمقصود بالوحدة التجريبية هنا هي أي حالة أو عنصر يمكن قياسه أو تقييمه بطريقة ما. والبيانات متعددة المتغيرات تظهر هنا متى ما قام الباحث بقياس أو تقييم أكثر من خاصية أو سمة واحدة لكل وحدة تجريبية. وهذه الخواص أو السمات تسمى عادة بالمتغيرات من قبل الإحصائيين.

طرق التحليل متعدد المتغيرات في غاية الأهمية لكونها تساعد الباحثين في تكوين إسنتاج فيما يخص مجتمع كبيرة ومتداخلة ومعقدة أحياناً من البيانات تتضمن عدداً كبيراً من المتغيرات مأخوذة من عدد كبير من الوحدات التجريبية. إن ضرورة وفائدة استخدام طرق التحليل متعدد المتغيرات تزداد بشكلٍ طردي مع زيادة عدد الوحدات التجريبية أو عدد المتغيرات المأخوذة عنها للتحليل.

وغالباً ما يكون الهدف من استخدام التحليل متعدد المتغيرات هو تلخيص الكمية الكبيرة من البيانات من خلال عدد صغير نسبياً من المعلمات Parameters. وبالتالي فإن الوظيفة الرئيسية لغالبية الأساليب متعددة المتغيرات هي التبسيط.

من جانب آخر، فإن التحليل متعدد المتغيرات غالباً ما يرتبط مع إيجاد علاقات ما بين:

- 1) متغيرات الإستجابة Response Variables
- 2) الوحدات التجريبية Experimental Units
- 3) كل من متغيرات الإستجابة والوحدات التجريبية

إن غالبية أساليب التحليل متعدد المتغيرات تمثل إلى كونها ذات طبيعة استكشافية لحالة ما بدلاً من كونها تأكيدية لتلك الحالة. وهذا يعني ميلها إلى خلق الفرضيات الإحصائية بدلاً من اختبارها.

وللتوسيح هذه النقطة، إفترض حالة كون باحثٍ ما لديه (50) خمسون متغيراً مقاسة على أكثر من (2000) ألفي وحدة تجريبية. إن الطرق الإحصائية الإعتيادية تتطلب من الباحث البدء بإدراج بعض الفرضيات أولاً ومن ثم قيامه بجمع البيانات، ومن بعد ذلك استخدام هذه البيانات لتبني الميل الإحتمالي لتبني (قول) هذه الفرضيات أو نفي ثبوتها صحتها (رفضها).

تصنيف طرق التحليل متعدد المتغيرات

أحد الفروقات الأساسية ما بين الطرق متعددة المتغيرات يكمن في تصنيفها إلى صنفين رئيسيين من حيث أساليب التحكم في التحليل وهما:

1) أساليب تحكم المتغيرات Variable – directed techniques

وتشمل تلك التي تتعامل بشكلٍ رئيسي مع العلاقات التي من الممكن ظهورها ضمن متغيرات الإستجابة Response Variables التي يتم قياسها. ومثال ذلك:

- التحليلات المعتمدة على مصفوفات معامل الإرتباط Correlation Matrices
- تحليل المركبات الرئيسية Principle Components Analysis
- التحليل العائلي Factor Analysis
- تحليل الإرتباط القوي (إرتباط المجاميع) Canonical Correlation Analysis

2) أساليب التحكم الشخصي Individual – directed techniques

وتشمل تلك التي تتعامل بشكلٍ رئيسي مع العلاقات التي من الممكن ظهورها ضمن الوحدات التجريبية و/أو الأشخاص الخاضعين للقياس. ومثال ذلك:

- التحليل المميز Discriminant Analysis
- التحليل العنقددي Cluster Analysis
- تحليل التباين المتعدد Multivariate Analysis of Variance (MANOVA)

عندما نبدأ التحليل

ما أن يبدأ الباحث التفكير في إجراء تحليل لمجموعة جديدة من البيانات، فإن أسئلة عديدة حول هذه البيانات يجب أن تؤخذ بالإعتبار. ومن هذه الأسئلة المهمة:

- 1) هل هناك أية جوانب في البيانات يمكن اعتبارها غريبة أو غير اعتيادية؟
- 2) هل يمكن الإفتراض بأن البيانات تتوزع طبيعياً Normality؟
- 3) هل هناك أي جانب منها خارج الطبيعي Abnormality؟
- 4) هل هناك أية قيم شاردة (شاذة) Outliers في البيانات؟ نقصد هنا بالقيمة الشاذة لوحدة التجربة حيث المتغيرات المقاسة منها تبدو وكأنها غير متناسبة مع القياسات المأخوذة لوحدات أخرى.

ولضرورة فهم التعامل مع المصفوفات والعمليات المرتبطة بها، نرى من الضروري البدء في تغطية هذا الموضوع قبل البدء في الموارد الرئيسية أعلاه.