الفصل

27

العلاج التجريبي

ظهرت في بداية الثمانينات، مشكلة لريتشارد بيتو Richard Peto، وهو أحد رواد الإحصاء الحيوي في بريطانيا العظمى. كان يقوم بتحليل نتائج التجارب الطبية، ويقارن بين أساليب علاج السرطان المختلفة بين المرضى، متبعاً تعليمات ر.آ. فيشر في تصميم التجربة. نختار للتجربة الطبية النموذجية مجموعة من المرضى ممن يحتاجون إلى العلاج، ثم نخضعهم عشوائياً، لأساليب مختلفة من التجارب العلاجية.

يجب أن تكون التحليلات واضحة نسبياً في مثل هذه البيانات. أُجريت المقارنة بين نسبة المرضى الذين قاوموا المرض لمدة خمس سنوات، وبين مجموعات العلاج الأخرى باستخدام أساليب فيشر. كان من الممكن إجراء مقارنة أخرى أكثر دقة، باستخدام اتجاه إيلان للمارتينغيل، لتحليل الزمن منذ

بدء الدراسة حتى وفاة كل مريض كمقياس أساسي لآثار المرض. كانت التحليلات في كلتا الحالتين مبنية على التوزيع العشوائي الأولي للمرضى في المعالجة. وحسب فيشر، لا يعتمد تخصيص علاج ما للمرضى كلياً على نتائج الدراسة، وبإمكاننا حساب قيم- p للاختبارات الفرضية.

كانت مشكلة بيتو في أنه لم يتم معالجة كل المرضى حسب توزيعهم العشوائي. فهؤلاء هم نفوس بشرية، تعاني من أمراض مؤلمة ومميتة في كثير من الحالات. شعر الطبيب المعالج بأنه مجبر على التنازل عن التجربة العلاجية، أو على الأقل على تعديلها إذا شعر أن ذلك لمصلحة المريض، إذ يكون الاتّباع الأعمى للمعالجة التحكّمية، من غير اعتبار حاجات المريض وردود أفعاله، أمراً غير أخلاقي. يتم تزويد المرضى، في مثل هذه الدراسات، بالعلاج الجديد في الوقت الذي يعتمد فيه اختيار العلاج تبعاً لاستجابة المريض، وذلك على عكس تعليمات فيشر.

كانت هذه مسألة نموذجية في دراسات مرض السرطان، وما زالت كذلك منذ أن ابتدأت هذه الدراسات في الخمسينات، إلى أن ظهر بيتو في الصورة. كانت الخطوات المتّبعة عبارة عن تحليل حالات المرضى الذين استمروا في المعالجة العشوائية فقط، وإلغاء كل المرضى الآخرين من التحليلات. لاحظ بيتو أن ذلك قد يؤدي إلى أخطاء جسيمة. فلنفترض مثلاً أننا نقارن بين معالجة فعالة وأخرى حيادية

Placebo، أي دواء حيوي لا تأثير له يعطى لإرضاء المريض. ولنفترض أن المرضى الذين لم يستجيبوا تم تحويلهم إلى العلاج التقليدي، وسيتم تحويل المرضى الذين فشل معهم العلاج الحيادي وإخراجهم من التحليلات . والمرضى الوحيدون الذين سيستمرون على العلاج الحيادي هم فقط المرضى، الذين استجابوا له، لسبب ما. ستظهر فعالية العلاج الحيادي (وربما أكثر) من العلاج الفعّال، إذا كان المرضى المستخدمون في التحليلات هم فقط الذين استمروا على العلاج الحيادي واستجابوا له.

لاحظ إدمند غيهن Edmund Gehan، من مستشفى م.س. أندرسون M.C.Anderson Hospital في تكساس، المسألة قبل بيتو. كان الحل الذي قدمه في ذلك الوقت بافتراض أن هذه الدراسات لم تحقق متطلبات فيشر، لذلك لا يمكن اعتبارها تجارب مفيدة لمقارنة العلاج. بدلاً من ذلك، فقد احتوت سجلات هذه الدراسات على ملاحظات دقيقة أخذت للمرضى الذين تم إعطاؤهم أنواعاً مختلفة من العلاج. إن أفضل ما يمكن توقعه هو وصف شامل لنتائجهم، مع تلميحات للمعالجة المستقبلية الممكنة. ثم قام غيهن بالنظر إلى حلول أخرى لهذه المسألة، ولكن كانت نتائجه الأولى تعكس انزعاج شخص يحاول تطبيق أساليب التحليلات الإحصائية لتجربة قديمة أو سيئة التصميم.

اقترح بيتو حلاً واضحاً. يتم اختيار المرضى عشوائياً

ليتلقوا علاجاً معيناً. جعلت هذه العشوائية حساب قيمة- p لاختبارات الفرضية لمقارنة أنواع المعالجة ممكنة. واقترح أن يُعامل كل مريض في هذه التحليلات، كما لو كان على العلاج الذي تم اختياره له عشوائياً. وسيهمل المحلل كل التغيرات العلاجية التي طرأت خلال فترة الدراسة. فإذا تم اختيار المريض عشوائياً على العلاج أ وتم إيقاف العلاج قبل إنهاء الدراسة سيحلل كمريض يعالج بالعلاج أ. وإذا تم اختيار المريض عشوائياً ليعالج بالعلاج أ واستمر على العلاج لمدة أسبوع، سيحلل كمريض يُعالج بالعلاج أ. وإذا لم يتناول مريض العلاج أ حبة واحدة من العلاج أ ولكن تم إدخاله إلى المستشفى وتمت معالجته بطرق أخرى فوراً بعد إدخاله في الدراسة، فإن هذا المريض سيحلل كمريض للعلاج أ.

قد يبدو هذا الاتجاه غريباً للوهلة الأولى. بإمكاننا أن نقدم رؤىً تتم المقارنة فيها بين العلاج القياسي والعلاج التجريبي، مع مرضى تم تحويلهم للعلاج القياسي عند فشل الأخرى. فإذا اتضح أن العلاج التجريبي غير مجدٍ، يتم تحويل كل المرضى أو معظمهم الذين تم اختيارهم للعلاج القياسي، وسيظهر في التحليلات تطابق العلاجين. فلا يمكن استخدام هذا الأسلوب في تحليل نتائج الدراسة لمعرفة تكافؤ العلاج، كما أوضح ريتشارد بيتو في عرضه. بل يمكن استخدامها فقط إذا ظهر اختلاف آثارها في التحليل.

أُطلق على حل بيتو أسلوب «الهدف من المعالجة». كان

سبب إطلاق هذا الاسم واستخداماته بشكل عام كالتالي: إذا كنا مهتمين بالنتائج الكلية للسياسة الطبية التي ترشح استخدام العلاج المُعطى، فيجب إعطاء الطبيب المُعالج الحرية في تعديل العلاج كما يراه ملائماً. ستقر تحليلات التجارب الطبية، باستخدام حل بيتو، فيما إذا كانت هذه السياسة العامة في اقتراح العلاج المُعطى كخطوة أولى في العلاج جيدة. تم تقديم تطبيقات أسلوب الهدف من المعالجة للتحليلات، كأسلوب مناسب للدراسات التي تتكفل فيها الحكومات الكبيرة، والمصممة لتحديد السياسات العامة الجيدة.

لسوء الحظ، هناك نزعة عند بعض العلماء في استعمال الأساليب الإحصائية من غير معرفة أو فهم المادة الرياضية خلفها. ويظهر هذا غالباً في عالم الأبحاث الطبية. لقد وضّح بيتو حدود الحل الذي قدّمه. ولكن أصبح، رغم ذلك، أسلوب الهدف من المعالجة مدخراً في الدستور الطبي في الكثير من الجامعات، وأصبح يُنظر إليه على أنه الطريقة الوحيدة والصحيحة للتحليلات الإحصائية للتجارب الطبية. صُمّمت الكثير من التجارب الطبية خاصةً المتعلقة بالسرطان، لتوضح أن العلاج الجديد هو بنفس جودة العلاج القياسي، أثناء عرض بعض الآثار الجانبية. إن الغرض من الكثير من التجارب هو توضيح التكافؤ العلاجي. وكما أشار بيتو، فإنه يمكن استخدام الحل المقدم للكشف فقط عن الفروق، ولا يعني عدم إيجاد فروق، أن أنواع العلاج متكافئة.

تكمن المشكلة، إلى حد ما، في تركيبة نيمان - بيرسون المتينة. يميل الاختصار المتعارف عليه في تركيبة نيمان - بيرسون الموجود في كتب الإحصاء الابتدائية إلى تقديم اختبار الفرضية كإجراء روتيني، بينما تعرض الكثير من الجوانب التحكّمية للأساليب على أنها ثابتة.

وبما أن الكثير من العناصر التحكّمية قد لا تكون مناسبة للبحث الطبي[1]، دعت حاجة بعض علماء الطب لاستعمال الأساليب «الصحيحة»، إلى الاحتفاظ بنسخة متينة من تركيبة نيمان - بيرسون. يُسمح بقبول أي شيء إلا إذا كانت قيمة- p

(1) قدم سنة 1963، فرانسيس أنسكومب Francis Anscombe من جامعة يل اتجاهاً مختلفاً تماماً يميل إلى الاستمرار مع الحاجات الطبية. تحتفظ تركيبة نيمان-بيرسون بعدد من الثمرات التي يخطئ فيها المحلل. يسأل أنسكومب لماذا لاحتمالية الخطأ الطويلة المدى للمحلل الإحصائي أن تقرر فيما إذا كان العلاج الطبي فعالاً. وكبديل لذلك، اقترح أنسكومب أن هناك عدداً نهائياً من المرضى سنتم معالجتهم. عدد قليل منهم سيعالجون بالتجارب الطبية. والآخرون سيتم إعطاؤهم العلاج الذي تقرره التجارب الطبية على أنه «الأفضل». وإذا استخدمنا عدداً قليلاً من المرضى في التجربة، فإن قرار اختيار العلاج الأفضل سيكون عرضة للخطأ، وإذا حصل ذلك، سيُعطى باقي المرضى العلاج الخاطئ. وإذا استخدمنا أعداداً كبيرة من المرضى في التجربة، فإن كل المرضى الخاضعين لتجربة العلاج الآخر (لا أعني العلاج «الأفضل») سيخضعون للعلاج الخاطئ. اقترح أنسكومب أن معيار التحليلات يجب أن يقلل من العدد الكلي للمرضى (لكل المرضى الخاضعين للعلاج التجريبي، والذين تمت معالجتهم لاحقاً) الذين أُخضعوا للعلاج الأضعف.

للطرق المختصرة ثابتة مسبقاً ومُحتفَظاً بها في الإجراءات الإحصائية. كان هذا أحد الأسباب التي عارض فيها فيشر تركيبة نيمان - بيرسون. لم يكن يعتقد وجوب خضوع استخدام قيم- p واختبارات الدلالة لمتطلبات دقيقة جداً. اعترض خصوصاً لقيام نيمان بتعديل احتمالية الموجب الوهمي مسبقاً، ولإتمامه عمله فقط إذا كانت قيمة- p أقل من ذلك. اقترح فيشر في كتابه الأساليب الإحصائية للاستدلالات العلمية، أن القرار الأخير حول دلالة قيمة-p يجب أن يعتمد على الظروف. استخدمت كلمة اقترح، لأنه لم يكن واضحاً لدى فيشر كيفية استعمال قيم- p. قام بتقديم الأمثلة فقط.

تركيبة كوكس

قام سنة 1977 ديفيد ر. كوكس (من بوكس وكوكس المذكورين في الفصل 23) بالتوسع في مناقشة فيشر. للتفريق بين استخدام فيشر لقيم- p وبين تركيبة نيمان-بيرسون، قام بتسمية أسلوب فيشر «اختبار الدلالة»، وتركيبة نيمان-بيرسون «اختبار الفرضية». أصبحت في الوقت الذي كتب فيه كوكس بحثه، حسابات الدلالة الإحصائية (من خلال استخدام قيم- p) من أكثر الأساليب استعمالاً في البحث العلمي. فاستنبط كوكس، أن الأسلوب له فائدته في العلوم. وبالرغم من النقاش الحاد بين فيشر ونيمان، وبالرغم من إصرار الإحصائيين مثل و. إدوارد ديمنغ بعدم فائدة الاختبارات الفرضية، وبالرغم من ظهور الإحصائيات البييزية التي لا وجود فيها لقيمة-p

وللدلالات- ورغم كل هذا النقد بين الإحصائيين الرياضيين، استمر استخدام اختبار الدلالة وقيم p. سأل كوكس، كيف يستخدم العلماء هذه الاختبارات؟ كيف بإمكانهم معرفة أن نتائج مثل هذه الاختبارات صحيحة أو مفيدة؟ اكتشف بالمران، أن العلماء يستخدمون اختبارات الفرضية في المقام الأول، لصقل نظرتهم لحقيقة الأمور، عن طريق التخلص من المتغيرات غير الضرورية، أو عن طريق الاختيار بين نموذجين مختلفين للحقيقة.

اتجاه بوكس

واجه جورج بوكس (النصف الآخر من بوكس وكوكس) المسألة من منظور مختلف قليلاً عن الآخر. أشار إلى أن الأبحاث العلمية تتألف من أكثر من تجربة واحدة. يُقدم العلماء على التجربة بقالب كبير من المعرفة المسبقة، أو على الأقل بتوقعات مسبقة عن النتائج الممكنة. ويتم تصميم الدراسة لتصقل هذه المعرفة، ويعتمد التصميم على نوع الصقل المنشود. يتفق على هذه النقطة كوكس مع بوكس في معظم أقوالهما. إن التجربة الواحدة، كما يراها بوكس، جزء من سيل من التجارب. وتتم مقارنة بيانات هذه التجربة ببيانات تجارب أخرى. ويُعاد النظر بالمعرفة السابقة تبعاً للتجربة الجديدة، وللتحليلات الجديدة للتجارب السابقة. لا يتوقف العلماء عن العودة للدراسات السابقة لصقل تفسيراتهم السابقة لها تبعاً للدراسات الجديدة.

وكمثال لاتجاه بوكس، فلننظر إلى صاحب مصنع الورق الذي يستخدم أحد ابتكارات بوكس الأساسية، التغير التطوري في العمليات EVOP. قدم صاحب المصنع التجارب، باستخدام ابتكار بوكس EVOP، إلى داخل العمل المنتج. تم تعديل درجة الرطوبة، والسرعة ونسبة الكبريت، ودرجة الحرارة قليلاً بعدة طرق. ولكن التغيير الناتج في قوة الورق ليس فائقاً. لا يمكن أن يكون فائقاً ويستمر في إنتاج المُنتج القابل للبيع. وبإخضاع هذه الاختلافات الطفيفة لتحليلات فيشر للتغاير، ممكن أن تستخدم في تقديم تجربة أخرى، يكون معدل قوة الورق فيها خلال العمل المُنتج بأكمله قد ازداد قليلاً، ويستخدم المُنتج الجديد للعثور على زيادة أخرى طفيفة في قوة الورق. تتم مقارنة نتائج كل مرحلة في EVOP بمراحل سابقة. تعاد التجارب التي تعطينا نتائج غير سوية ثانية. ويستمر هذا الإجراء لفترة طويلة، لا يوجد هناك حل نهائي «صحيح». لا يوجد في نموذج بوكس، نهاية لسلسلة التجارب العلمية التي تتبعها الفحوص المتكررة للبيانات، لا توجد هناك حقيقة علمية نهائية.

رؤية ديمنغ

رفض ديمنغ وغيره من الإحصائيين استخدام اختبارات الفرضية كلية. أصروا على أن أعمال فيشر في أساليب الحساب يجب أن تشكل قاعدة من التحليلات الإحصائية. إن متغيرات التوزيع هي التي يجب حسابها. ومن غير المجدي إجراء

تحليلات تتعامل بشكل غير مباشر مع هذه المتغيرات من خلال قيم- p والفرضيات الاعتباطية. استمر هؤلاء الإحصائيون باستخدام فترات الثقة لنيمان لقياس عدم دقة نتائجهم؛ ولكن أصبح اختبار الفرضيات لنيمان-بيرسون كما صرحوا، في نهايات الماضي، مع أسلوب كارل بيرسون في العزم. من الممتع أن نشير أن نيمان نفسه، كان نادراً ما يستخدم قيم- p واختبارات الفرضية في أبحاثه التطبيقية.

يمكن لهذا الرفض تجاه اختبار الفرضيات، وإعادة تركيب بوكس وكوكس لمفهوم فيشر لاختبار الدلالة، أن يطرح شكوكاً حول حل ريتشارد بيتو للمسألة التي وجدها في الدراسات الطبية للسرطان. ولكن ما زالت المسألة الأساسية التي واجهها قائمة. ماذا تفعل عندما يتم تعديل التجربة بالسماح لنتائج المعالجة بتغيير العلاج؟ أوضح أبراهام والد كيف يمكن تكييف نوع معين من التعديلات، لتؤدي إلى تحليلات متوالية. في حالة بيتو، لم يتبع أطباء الأورام أساليب والد المتوالية، بل أدخلوا معالجات مختلفة بحسب تقديرهم للحاجة.

دراسات كوشران المتعلقة بالمشاهدة

عالج وليام كوشران من جامعة جونز هوبكينز Johns Hopkins University، بطريقة ما هذه المسألة في الستينيات. كان على مدينة بالتمور أن تقرر، أثر الإسكان العام على السلوك الاجتماعي وعلى تطوير الفقراء. توجهوا إلى مجموعة جونز

هوبكينز الإحصائية لمساعدتهم في إعداد تجربة ما. اقترح الإحصائيون في جامعة جونز هوبكينز، متبعين أساليب فيشر، أن يأخذوا مجموعة من الأشخاص، سواء كانوا قد قدموا طلباً إلى الإسكان العام أو لم يفعلوا، وقاموا باختيار بعضهم عشوائياً للمساكن الخاصة ورفضوا الآخرين. أخاف هذا المسؤولين في المدينة. كان عليهم عند بدء التسجيل في الإسكان العام، الاستجابة على أساس الخدمة للمتقدم الأول. كان هذا منصفاً. لم يستطيعوا إنكار واجبهم تجاه الأشخاص الذين أسرعوا في التسجيل ليكونوا من «الأوائل»، وكان هذا على أساس الاختيار العشوائي للكمبيوتر. أشارت مجموعة إحصائيي جامعة جونز هوبكينز، أن هؤلاء الذين سارعوا في التقديم كانوا من الأشخاص الأكثر نشاطاً وطموحاً. إذا افترضنا صحة ذلك، سيكون الأشخاص في الإسكان العام أفضل حالاً من غيرهم، من غير أي تأثير للمساكن بحد ذاتها على ذلك.

اقترح كوشران أنهم لن يتمكنوا من استخدام تجربة مصممة علمياً. وستتم عوضاً عن ذلك، متابعة العائلات التي اشتركت بالإسكان العام. سيكون للتي لم تشترك دراسة مبنية على المشاهدة، تختلف فيها العائلات بعدة عوامل، مثل العمر والمستوى التعليمي والديانة والاستقرار العائلي. اقترح أساليب لإجراء التحليلات الإحصائية لمثل هذه الدراسات المبنية على المشاهدة. سيقوم بتنفيذ ذلك عن طريق تعديل المقاييس الناتجة للعائلة المفترضة، وبأن يأخذ بالحسبان العوامل المختلفة.

سيقوم بإنشاء نموذج رياضي يكون فيه تأثير لعامل العمر، وفي كون العائلة سليمة، ولعامل الدين، وهكذا دواليك. يمكن استخدام الفروق المتبقية في تحديد أثر الإسكان العام بحساب المتغيرات لكل هذه الآثار.

وعندما تعلن الدراسة الطبية أن الاختلاف في الآثار تم تعديله تبعاً لعمر المريض أو جنسه، يعني هذا أن الباحثين قاموا بتطبيق بعض أساليب كوشران في حساب الأثر الضمني للعلاج، واضعين في الحسبان أثر عدم التوازن في مهمة العلاج بالنسبة للمرضى. تستخدم معظم دراسات علم الاجتماع أساليب كوشران. قد لا يلاحظ مؤلفو هذه الدراسات أن هذه الأساليب مقتبسة من وليم كوشران، وكثير من التقنيات الخاصة غالباً ما سبقت أعماله. وضعها كوشران على أساس رياضي متين، أثرت أبحاثه حول الدراسات المبنية على المشاهدة في الطب، وعلم الاجتماع، والعلوم السياسية، وعلم الفضاء، وكل المجالات التي تكون فيها المهمات العشوائية في «المعالجة» إما مستحيلة أو لاأخلاقية.

نماذج روبن

تقدم دونالد روبن Donald Rubin من جامعة هارفرد، في الثمانينيات والتسعينيات، بعدة اتجاهات لمسألة بيتو. من المفترض في نموذج روبن، أن يكون لكل مريض استجابات ممكنة تجاه كلا العلاجين. فإذا وُجد علاجان، سيكون لكل

مريض استجابات احتمالية لكلا العلاجين أ وب. يمكننا مراقبة المريض عند تناوله أحد العلاجين، العلاج الذي تم تحديده له. بإمكاننا أن نعد نماذج رياضية يكون فيها رمز في التركيبة لكل الاستجابات الممكنة .قام روبن باشتقاق شروط في هذا النموذج الرياضي، قد نحتاجها لحساب ما يمكن حدوثه في حال وضعنا المريض على العلاج الآخر.

بإمكاننا تطبيق نماذج روبن وأساليب كوشران على التحليلات الإحصائية الحديثة، لأنها تستخدم الكمبيوتر في طحن الأرقام والتعامل معها جيداً. حتى لو تم تقديمها في زمن فيشر، لأنها لن تكون معقولة. إنها تتطلب استخدام الكمبيوتر لأن النماذج الرياضية متداخلة ومعقدة بشكل هائل. وغالباً ما تتطلب تكنولوجيا تكرارية، ونجد أن الكمبيوتر يقوم بآلاف بل بملايين الحسابات، تتقارب فيها سلسلة الحسابات المتتالية في الإجابة النهائية.

تعتبر أساليب كوشران وروبن نماذج عالية المستوى. أي أنها لا تنتج إجابات صحيحة، إلا إذا تقاربت النماذج الرياضية المعقدة المستخدمة في وصف الحقيقة. يطلب من المحلل أن يبتكر نموذجاً رياضياً يتطابق مع الحقيقة في كل أو في بعض الجوانب. وقد لا تتماسك نتائج التحليلات إذا لم تتطابق الحقيقة مع النموذج. وكجزء مصاحب للاتجاهات مثل اتجاه كوشران وروبن كان الجهد في تحديد درجة متانة النتائج. تبحث التحريات الرياضية الحالية عن مدى بعد الحقيقة عن

النموذج قبل إثبات عدم صحة النتائج. تفحص وليم كوشران هذه الأسئلة قبل موته سنة 1980.

يمكننا اعتبار أساليب التحليلات الإحصائية وكأنها سلسلة متصلة، ذات رابطة نموذجية عالية مثل التي قدمها كوشران وروبن في أحد الأطراف. توجد في الطرف الآخر، الأساليب التي لا تستخدم المتغيرات، والتي تفحص البيانات حسب النماذج ذات النوع الأكثر عمومية. ومثل ما جعل الكمبيوتر الأساليب ذات الرابطة النموذجية ممكنة، لقد كانت ثورة الكمبيوتر في الطرف الآخر من تشكيل النماذج الإحصائية. وهذه النهاية التي لا متغيرات لها، والتي يفترض أن يكون لها بنية رياضية ضعيفة أو معدومة، والتي يسمح فيها للبيانات بذكر كل ما فيها من غير أن تخضعها لنماذج سبق التفكير فيها. وتصاحب هذه الأساليب أسماء وهمية مثل «سبر الحذاء». وهو موضوع حديثنا في الفصل التالي.