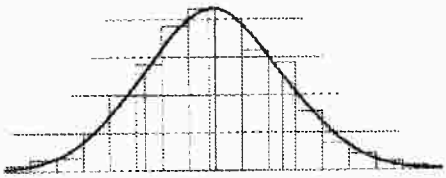

الفصـــــل

22

بيكاسو عصره في الإحصاء

أجريت بعد ما أنهيت رسالة الدكتوراه سنة 1966، جولات على عدة جامعات، وتحدثت فيها عن نتائج أبحاثي كما أجريت مقابلات للحصول على عمل. كانت جامعة برينستون من أوائل المحطات التي توقفت عندها، وتقابلت هناك في محطة القطار مع جون تاكي.

سمعت خلال دراستي عن تناقضات تاكي، ودرجة الحرية في تفاعلاته، وعن تحويلات فورييه Fourrier، السريعة له، وكذلك عن اختباره السريع، وفرضيته المساعدة ما يسمى بالتوطئة. ولم أدرس إلى الآن أعماله في تحليل البيانات الاستكشافي، أو أي من الأعمال التي صدرت من عقله المبدع في السنين اللاحقة. كان تاكي رئيس اللجنة في قسم الإحصاء، (وكان لديه منصب مشترك في مختبرات شركة الهاتف بيل Bell

Telephone Laboratories)، وقد فاجأني عندما قام هو شخصياً ليرحب بوصولي. كان يرتدي البنطلون الصيني مع قميص قطني واسع فوقه، وحذاء رياضي. بينما كنت أرتدي بذلة وربطة عنق. لم تكن ثورة الستينات للملابس المخيطة قد وصلت إلى مستوى الكلية، لذلك بدت ملابسي ملائمة أكثر من ملابسه.

تمشينا سوية حول السكن الجامعي، نناقش أحوال المعيشة في برينستون. سألني عن برامج الكمبيوتر التي كتبتها أثناء العمل برسالتي. وضّح لي بعض الحيل لتجنب خطأ التدوير (الأخطاء المتراكمة) في برامجي. وصلنا أخيراً إلى القاعة التي سأتحدث فيها عن رسالتي. صعد تاكي إلى الصف الأخير من المقاعد في الغرفة بعد تقديمي للحاضرين. بدأت بتقديم النتائج التي توصلت إليها. لاحظت انشغاله بتصحيح أوراقه أثناء ذلك.

سألني بعدها بعض الحاضرين (الذين كان معظمهم من الخريجين وأعضاء الكلية) أسئلة، واقترح آخرون بعض التشعبات لما قدمت. ثم تقدم جون تاكي من الصف الأخير عند انتهاء الجميع من أسئلتهم وتعليقاتهم. أخذ قطعة طباشير وأعاد تقديم نظريتي الأساسية على اللوح المدرسي، واضعاً كل الرموز الرياضية التي استخدمتها[1]. ثم عرض علي برهاناً آخر

(1) أخافت الرموز الرياضية المتكونة من مجموعة من الحروف، الرومانية واليونانية، مع خطوط غير متقنة، ورموز وحروف سفلية وعلوية، الكثير من غير الرياضيين (وغالباً بعض الرياضيين أيضاً). فهي طريقة ملائمة =

لنظرية استغرق إثباتها بطريقتي الخاصة شهوراً عدة. «ما أزوع ذلك»، قلت في نفسي. «هذا تحالف عظيم!».

ولد جون تاكي سنة 1915 في نيوبدفورد ماساشوسيتس New Bedford Massachusetts. كان يخالط كلامه اللهجة القريبة من أهل بوسطن. لاحظ والداه عبقريته مبكراً، وتلقى العلم في المنزل إلى أن غادر إلى جامعة براون Brown University، فحصل هناك على درجة الماجستير في الكيمياء. أثارت الرياضيات المجردة اهتمامه. أتم دراسته العليا في جامعة برينستون، وهناك حصل على شهادة الدكتوراه في الرياضيات سنة 1939. كانت أعماله الأولى في مجال الطوبولوجيا. تعتبر

= لربط الأفكار المعقدة بعضها ببعض في الفضاء المدمج. «والفن» في قراءة البحث الرياضي، هو إدراك أن كل رمز له معنى خاص، نعرفه عند تقديمه، ولكن بعدها يجب أن نثق بأننا «فهمنا» معناه، ونركز انتباهنا على الطريقة التي يُستخدم فيها الرمز. يكمن روح الذوق الرياضي في تقديم مجموعة من الرموز، يسهل للقارئ ملاحظتها وفهم العلاقات بينها بسرعة. ممكن أن نجد هذا النوع من الذوق الرياضي في أبحاث جيرزي نيمان. كانت رسالتي للدكتوراه بعيدة عن الذوق على ما يبدو. كنت استخدم الرموز الرياضية للتأكد من أن شمول جميع الجوانب الممكنة للأساليب الرياضية. كان للرموز والحروف السفلية رموز أخرى مثلها والرموز العليا كان لها أيضاً رموز مثلها وفي بعض الحالات، كان للرموز السفلية عدة أوجه مختلفة. أذهلني جون تاكي عندما استطاع إعادة صياغة هذا الكم المعقد من الرموز، لأول مرة كان يرى فيها النظرية في ظهيرة ذلك اليوم. (بالرغم من الفوضى السائدة بين رموزي، فقد عرض علي تاكي عملاً. ولكن كان لدي ثلاثة أطفال وآخر قادم في طريقه، قبلت عملاً آخر بأجر أكثر في مكان آخر).

دعامة الطوبولوجيا دعامة أساسية ونظرية في الرياضيات. ودون دعامة الطوبولوجيا فرع صعب وسري من علم المنطق، يعرف بـ«ما وراء الرياضيات». يتعامل ما وراء الرياضيات مع أسئلة عمّا يعنيه «حل» مسألة رياضية، وماهية الفرضيات غير المصرح بها وراء استخدام المنطق. تعمق تاكي في هذه الدعائم المظلمة واستخرج منها «الفرضية المساعدة لتاكي»، التي تعتبر إنجازاً أساسياً في هذا المجال.

لم يكن مقدراً لجون تاكي البقاء في مجال الرياضيات المجردة، بينما كان ساموبل س. ويلكس في كلية الرياضيات في برينستون، يدفع كثيراً من الطلبة وأعضاء الكلية من الشباب للالتحاق بعالم الإحصاء الرياضي. أطلق عليه لقب معلم في قسم الرياضيات في الجامعة حين حصوله على الدكتوراه. وفي سنة 1938، وأثناء عمله بالرسالة، قام بنشر بحثه الأول حول الإحصاء الرياضي. وكانت معظم منشوراته بعد سنة 1944 في هذا المجال.

التحق تاكي بمكتب أبحاث مراقبة الحرائق خلال الحرب العالمية الثانية، ليعمل على حل مشاكل التصويب في البنادق، وتقييم إحكام الرمي في الآلات الحربية، ومشاكل أخرى متعلقة بالمعدات الحربية. زودته هذه التجربة بالكثير من الأمثلة للمسائل الإحصائية التي سيحقق فيها مستقبلاً، وجعلته يقدر طبيعة المسائل العملية. اشتهر بحِكَمِه المؤلفة من سطر واحد التي شملت تجارب هامة. وهذه مثلاً حكمة من تجربته العملية

«من الأفضل أن يكون لدينا إجابة تقريبية للسؤال الصحيح، أكثر من أن يكون لدينا إجابة دقيقة للسؤال الخطأ».

تعدد جوانب تاكي

أذهل بابلو بيكاسو Pablo Picaso عالم الفن في بداية القرن العشرين بمنتجاته المختلفة في روعتها. كان يتفنن باللوحات الزيتية أحادية اللون، ثم اخترع التكعيبية، وبعدها قام بتجريب نوع من الكلاسيكية، وبعدها اتجه إلى الخزف. نتج عن هذه التقلبات ثورات متغيرة في الفن، والتي استمر في استثمارها آخرون بعد اتجاه بيكاسو لأمور أخرى. هكذا كان الأمر مع تاكي: في الخمسينات أصبح مهتماً بأفكار أندريه كولموغوروف حول مراحل الاتفاقية، وطور كمبيوتراً له تقنية خاصة لتحليل العناصر المتتالية الطويلة للنتائج المرتبطة ببعضها. عرفت «بتحويلات فورييه السريعة Fast Fourier transfom»، أو FFT. سيبقى تأثير تاكي على العلوم هائلاً مثل بيكاسو والتكعيبية، حتى لو توقفت إنجازاته عند هذا الحد.

كان عمل تاكي في الحرب سبباً في نقله سنة 1945 إلى مختبرات شركة الهاتف بل Bell في ماري هيل، نيوجيرسي Murray Hill، New Jersey، حيث انهمك في المسائل العملية المختلفة. «كان لدينا مهندس اسمه بادنبوم Budenbom»، قالها في أحد المحادثات المسجلة سنة 1987، يقوم ببناء رادار جديد ممتاز لتعقب الآثر من أجل اقتفاء أثر الأهداف الجوية. كان يريد الذهاب إلى كاليفورنيا ليقدم بحثه، وطلب صورة ليوضح لهم عن

نوعية أخطاء تعقب الأثر خاصته». قام بادنبوم بكتابة مسألته بنطاق التردد، ولكنه لم يعرف كيف يستخرج تقديرات متماسكة لسعة التردد. لم يتعرض تاكي بعد لاستخدامات هذه التقنية في الهندسة، رغم أنه كان عالم رياضيات، وكان على علم كاف بتحويلات فورييه. قدم أسلوباً أعجب المهندس (فلنذكر حكمة تاكي عن فائدة الإجابة التقريبية للسؤال الصحيح). لكن تاكي نفسه لم يكن مقتنعاً، فاستمر بالتفكير في المسألة.

كانت النتيجة تحويل فورييه السريع. وهي تقنية مصقولة باستخدام تعبير تاكي، «تستلف القوة» من الترددات المجاورة، وكمية كبيرة من البيانات غير لازمة لتقديرات جيدة. وهي أيضاً حل نظري تم التفكير به جيداً مع صفات مثلى. هو كمبيوتر في نهاية الأمر ذو كفاءة عالية في الحساب. كانوا بحاجة لمثل هذه الحسابات في الخمسينيات والستينيات، عندما كان الكمبيوتر بطيئاً وذا ذاكرة محدودة. استمر استخدامها في القرن الحادي والعشرين، لأنها متفوقة في دقتها لبعض التقديرات للتحويلات المعقدة التي أصبح يمكن تنفيذها اليوم.

استمر تقدم الكمبيوتر وطاقاته في الحقول الجديدة للبحث الإحصائي. رأينا كيف جعل الكمبيوتر حساب المصفوفات الكبيرة المعكوسة ممكناً، (وهذا الأمر قد يستغرق من جيري كورنفيلد عمل مئات السنوات على آلته الحاسبة الميكانيكية). هناك جانب آخر في الكمبيوتر يهدد بسحق النظرية الإحصائية، وهو قدرته على تخزين وتحليل كمية كبيرة من البيانات.

تفوق مهندسو وإحصائيو مختبرات شركة الهاتف بل في الستينات وبداية السبعينات، في التحليلات التي أجريت على الكميات الهائلة من البيانات. أدت مراقبة خطوط الهاتف والأخطاء العشوائية والمسائل، إلى ملايين البيانات للملف الواحد في الكمبيوتر. وأُرسلت البيانات التي قيست أبعادها من مجس فضائي، لفحص كوكب عطارد والمريخ وغيرهما من الكواكب، فأنتجت ملفات أخرى من ملايين المفردات. كيف يمكنك النظر إلى مثل هذه البيانات؟ وكيف تبدأ بتنظيمها لتصبح جاهزة للفحص؟

بإمكاننا دوماً حساب المتغيرات للتوزيعات الاحتمالية، متبعين تقنية الرائد كارل بيرسون. يتطلب منا هذا أن نفترض أمراً عن هذه المتغيرات - إنها تتبع نظام بيرسون، على سبيل المثال. هل يمكننا أن نجد أساليب لاختبار هذه المجموعات الضخمة من الأرقام، ونتعلم شيئاً منها من غير أن نفرض افتراضات عن توزيعاتها؟ هذا ما كان يفعله نوعاً ما العلماء الجيدون. أجرى غريغور ميندل Gregor Mendel سلسلة من التجارب على تهجين النبات، فاحصاً المحصول ومطوراً بالتدريج نظرياته للجينات المسيطرة والمتنحية. بالرغم من أن كثيراً من الأبحاث العلمية تضمنت جمع البيانات وملاءمة هذه البيانات للقالب المعد سابقاً لنوع خاص من التوزيعات، غالباً ما قد يكون مفيداً وهاماً جمع البيانات وفحصها بدقة للحوادث غير المتوقعة.

وكما أشار مرةً أحد علماء الرياضيات الأمريكيين إيريك تمبل بل Eric Temple Bell، «لا تكذب الأرقام، ولكن لديها النزعة لقول الحقيقة مع نية للخداع»[2]. تميل النفس البشرية لرؤية الأمثلة وغالباً ما سترى الأمثلة فقط مع الضجة العشوائية[3].

أزعج هذا الأمر علماء الأوبئة على وجه الخصوص، فغالباً ما كانوا يجدون «عناقيد» من الأمراض، في أوقات وأماكن محددة أثناء فحصهم مجموعات من البيانات. فلنفترض أننا وجدنا عدداً كبيراً من الأطفال المصابين بسرطان الدم في مدينة صغيرة في ماساتشوستس Massachusetts، فهل يعني ذلك وجود سبب معين لانتشار هذا المرض في هذه المدينة؟ أم أنها مجرد عناقيد عشوائية ظهرت الآن

(2) كتب بل عدة كتب مشهورة في الرياضيات في الأربعينات والخمسينات. مازالت سيرة قادته من علماء الرياضيات، المرجع التقليدي لرواد الرياضيات في القرن الثامن والتاسع عشر. كانت دراسته لمعاني الأعداد، والتي أخذنا منها هذه المقولة، تخبر عن معاني الأعداد، والتي عرضت عليه، كما كتب، من السيدة التي كانت تتظف له.

(3) وكمثال كلاسيكي لذلك هو قانون بود بود Bode's law. وقانون بود هو المشاهدة المعتمدة على التجربة، وهي بوجود علاقة خطية بين لوغاريتم المسافة من الشمس، وعدد الكواكب في النظام الشمسي. تم اكتشاف نبتون لأن علماء الفضاء طبقوا قانون بود لتوقع المدار التقريبي لكواكب أخرى ووجدوا نبتون في هذا المدار. ولحين تحقق علماء الفضاء من وجود زحل والمشتري، اكتشفوا الكثير من الأقمار الصغيرة بالقرب من الكواكب، واتضح أن القمر الصناعي الوحيد الذي استطاع أن يشاهد زحل كان يتبع قانون بود. هل قانون بود هو عبارة عن مصادفات عشوائية؟ أو هل يقوم بإخبارنا بأمور عميقة وبالتالي غير مفهومة عن العلاقات بين الكواكب والشمس؟

في هذا المكان، وكان ممكناً ظهورها ببساطة في مكان آخر؟ ولنفترض أن سكان المدينة اكتشفوا مثلاً أن مصنعاً كيماوياً يتخلص من نفاياته في بحيرة قريبة، أو أنهم اكتشفوا آثاراً لأحماض أمينية في موارد مياه المدينة التي ظهرت فيها عناقيد مرض سرطان الدم. فهل يكون ذلك سبباً للإصابة بالمرض؟ أو بعبارة أشمل، إلى أي مدى يمكننا فحص البيانات متتبعين الأنماط والظواهر، ومتوقعين الحصول على نتائج غير السراب العشوائي؟

بدأ جون تاكي في الستينات، بالبحث جاداً في هذه المسائل، واستخرج بذلك نسخة منقحة لاتجاه كارل بيرسون نحو البيانات. لاحظ إمكانية فحص توزيع البيانات الملحوظة كتوزيع، من غير فرض نموذج احتمالي اعتباطي عليه. كانت النتيجة سلسلة من الأبحاث ومحادثات في الاجتماعات، فكتب أخيرا كتباً تُبحث فيها أسماء «تحليلات استكشافية للبيانات». وأثناء عمله بالمسائل، كان لنمط تاكي في التقديم شكل أخّاذ ولافت للنظر. بدأ بإعطاء أسماء مختلفة للصفات المميزة للتوزيعات البيانية، غير التي كانت تُستخدم في الماضي، وذلك ليصدم مستمعيه وقراءه نحو إعادة فحص فرضياتهم. كما أنه ابتعد عن التوزيعات الاحتمالية المعتادة، مثل نقطة بداية التحليلات إلى فحص الأمثلة ضمن البيانات نفسها. لقد كان ينظر إلى الطريقة التي تؤثر فيها القيم المبالغ فيها على مراقبتنا للأمثلة. ولتعديل مثل هذه التأثيرات الكاذبة، قام بتطوير مجموعة من الأدوات البيانية لعرض البيانات.

بيّن على سبيل المثال، أن المدرج التكراري المعروف لعرض توزيع مؤلف من مجموعة من البيانات، قد يكون مضللاً لتوجيه النظر لمتابعة مجموعة المشاهدات الأكثر تردداً. افترض أنه بإمكاننا تعيين الجذر التربيعي للتردد عوضاً عن تعيين تردد المشاهدات. وأطلق على ذلك «الجذر التكراري» موازنة بالمدرج التكراري. افترض تاكي أيضاً أنه يمكن تعيين الحقل الرئيسي للبيانات كصندوق، وتعيين القيم القصوى بخطوط (أو كما أطلق عليها «بالشعيرات») تنبعث من الصندوق. أصبحت بعض الأدوات التي قدمها جزءاً لكثير من البرامج الإحصائية المعروفة، وأصبح بإمكان المحلل أن يطلب «مخطط الصندوق» «ومخطط الساق والورقة». لقد انجرفت مخيلة تاكي الخصبة خلال تحليلات الصورة البيانية، وتندمج كثيراً من عروضه مع برامج الكمبيوتر. تم دمج اثنين من اكتشافاته للغة الإنجليزية. فهو من ابتكر كلمة bit لتعني («binary digit» «رقم ثنائي») وكلمة سوفتوير software (برامج الكمبيوتر، تماشياً مع هاردوير «hardware» التي تعني الكمبيوتر نفسه).

لم يصعب على تاكي أن يهاجم الأمور برؤيته البديهية، ولم يكن هناك شيء مقدس بالنسبة له ليسأل عنه. فلنأخذ على سبيل المثال مرحلة التسجيل البسيطة. تعرّض معظم القراء لاستخدام عصا الحساب أثناء عد أي شيء. إنها العصا المعتادة التي قدّمها أجيال من الأساتذة، فنضع أربع علامات عمودية والخامسة لتشطب الأربع. كم من أفلام الكرتون شاهد فيها

القارئ، ذاك السجين الأشعث وهو يخطّ سلسلة من هذه الخطوط بالطباشير على حائط السجن؟

يقول تاكي إن استخدام العصي طريقة حمقاء. فلننظر إلى سهولة الوقوع في الخطأ. فقد تضع العلامة لتشطب ثلاثة خطوط بدلاً من أربعة، أو قد تضع خمسة خطوط ثم تشطبها. من الصعب تحديد الخطأ بهذه الطريقة. فقد تبدو صحيحة إلا إذا دققت في عدد الخطوط العمودية. قد يبدو معقولاً أن نستخدم العصا في الأماكن التي يسهل فيها تحديد الخطأ. عرض تاكي عشرة خطوط أو عصي أمامنا. نقوم في البداية برسم أربع نقاط لتحديد زوايا الصندوق. ثم نصل بين النقاط بأربعة خطوط لننهي الشكل. وفي النهاية نرسم خطين قُطرين أو مائلين لتشكل علامة تقاطع داخل الصندوق.

كانت هذه الأمثلة من تحويلات فورييه السريعة، وتحليلات البيانات الاستكشافية، جزءاً من نتاج تاكي الضخم. مثله مثل بيكاسو وهو يتنقل من الرسم التكعيبي إلى الكلاسيكي، ثم إلى الرسم على الخزف وعلى الأقمشة. تنقل جون تاكي عبر الصورة الإحصائية الجميلة للنصف الثاني من القرن العشرين، من السلسلة الزمنية، إلى النماذج الخطية، ثم إلى تعميم بعض من أعمال فيشر المنسية، وإلى تنشيط الحسابات، فتحليلات البيانات الاستكشافية. أنشأ من النظرية المتأصلة في الرياضيات، اعتباراً للمسائل العملية، وأخيراً وضع اعتباراً للتقييم غير المنظم للبيانات. تغير مفهوم الإحصاء كلما

وضع يده فيه. كان جون تاكي، وحتى يوم وفاته في صيف 2000، مستمراً في تحدي زملائه ومساعديه بالأفكار الجديدة والتساؤل المستمر عن القديم.