

الفصل الخامس

المبادئ الإحصائية وعمليات الفرز السريع

Chapter Five

Statistical Principles and Quick Counts

تُطبّق منهجية الفرز السريع المبادئ الإحصائية على مشكلة عملية للغاية، ألا وهي التحقق من النتائج الانتخابية.¹ يوجز هذا الفصل هذه المبادئ الإحصائية ويصف كيفية عملها معاً. أمّا الطريقة الأكثر اختصاراً في إظهار هذه المعلومات فتكمن في استعمال لغة الرياضيات وهي طريقة لا مناص منها إلى حد ما. يرمي هذا الفصل إلى تقديم مفاهيم أساسية، على نحو غير فني، لإتاحة اطلاع الجمهور على المنطق الذي تركز عليه منهجية الفرز السريع.

يقدم الجزء الأول من هذا الفصل أسس منهجية الفرز السريع. يبدأ بالتأكيد على صلاية معطيات الفرز السريع، وعلى مفاهيم جوهرية كالعينة والسكان. ثم يتحول الفصل إلى شرح المبادئ الإحصائية كقانون الأعداد الكبيرة ونظرية الحد المركزي. أما النصف الثاني من الفصل الذي يتسم بطابع فني أكثر فيقدم عملية أخذ العينات. وما يلبث أن يعرف إجراءات النزعة المركزية والتشتت ثم يناقش استراتيجيات معيارية لاحتساب العينات واستخراجها. يتطرق أيضاً إلى مسائل عملية كعوامل التصحيح المخصصة لإدارة المشاكل الفريدة التي تنشأ لدى تطبيق المبادئ الإحصائية في حالات الفرز السريع.

المبادئ الإحصائية الأساسية

تتحمّك المبادئ الإحصائية بمنهجية جمع معطيات الفرز السريع وتحليلها. تركز هذه المنهجية على مبادئ علمية مسلّم بها إلى حد كبير. على غرار قانون الجاذبية، ليست هذه المبادئ الإحصائية مجرد مسألة رأي كما أنها

¹ يركز هذا الفصل على المبادئ الإحصائية، المستخدمة في أخذ عينة عشوائية من مراكز الاقتراع التي تُجمع منها المعطيات وتحلل، من أجل التحقق من النتائج الانتخابية. غير أن منهجية الفرز السريع تطوّرت، والمبادئ الإحصائية هي ذاتها التي تدير حالياً مراقبة الانتخابات النوعية. يصف الفصل السادس، العنصر النوعي في الفرز السريع، كيفية جمع المعلومات حول عمليات التصويت وفرز الأصوات من المراقبين عينهم، ومراكز الاقتراع ذاتها التي تستعمل لاستخراج المعطيات حول فرز الأصوات. يمكن تعميم هذه النتائج بشكل موثوق به على نوعية عمليات التصويت والفرز في أرجاء البلاد.

ليست مفتوحة على التأويل الحزبي؛ من هنا، يمكن إثباتها وتقبلها على صعيد عالمي. وبفضل ارتكاز هذه المبادئ على أساس علمي، بإمكان منظمي الفرز السريع أن يقدموا ادعاءات جازمة حول نتائج الانتخابات. لا يتعلق الأمر بتحديد عدل الانتخابات أم ظلمها تماماً؛ فمنهجية الفرز السريع تتيح للمجموعة أن تثبت سبب اعتبار عمليات اليوم الانتخابي عادلة، أو إلى أي حد لم تكن عادلة.

تتيح منهجية الفرز السريع للمجموعة أن

تثبت سبب اعتبار

عمليات اليوم

الانتخابي عادلة أو

إلى أي حد لم تكن

عادلة.

الموثوقية والشرعية

لا تعتبر البيانات الصادرة حول عمليات اليوم الانتخابي قوية إلا بقدر قوة المعطيات التي ترتكز عليها. وبناء عليه، لا بد من خطو خطوات ثابتة ومدروسة لضمان استيفاء المعطيات المجموعة بعض المعايير. تتمثل هذه الخطوات في تمتع معطيات الفرز السريع، بحد ذاتها، "بالقوة". ونقصد بهذا المفهوم أن تكون موثوقاً بها وصالحة، أي مسلم بصحتها.

تعتبر المعطيات موثوقاً بها عندما يقيم المراقبون المسقلون الذين يراقبون الحدث ذاته (الفرز السريع) ويستعملون أدوات القياس نفسها (استمارة المراقب) الحدث بالطريقة نفسها. يظهر المثل التالي ما نعبه:

يقيس ثلاثة أشخاص (أ-ب-ج) بصورة متكررة طول الشخص الرابع (ي) في اليوم نفسه. يعتبر قياس طول هذا الشخص موثوقاً به في حال توصل المراقبون الثلاثة (أ-ب-ج) الذين يستعملون أدوات القياس ذاتها (شريط قياس عادي) إلى النتائج نفسها، على نحو دقيق، في ما يخص طول الشخص الرابع.

ينطبق هذا المبدأ ذاته على جمع معطيات الفرز السريع؛ فموثوقية المؤشرات والقياسات تشكل أمراً أساسياً. يجب ألا تتبدل المعلومات التي تصدر عن المراقبين بسبب المؤشرات غير الكافية أو أدوات القياس غير الملائمة (شريط قياس مطاطي) أو التداوير الرديئة—ولا يجب أن تتفاوت النتائج بحسب الشخص الذي يجري القياس. لا تتفاوت النتائج الموثوق بها إلا عند حصول تبدلات حقيقية في الظاهرة التي يجري قياسها. وبالتالي، تُعد المعطيات التي يمكن الوثوق بها المعلومات التي يمكن التحقق منها بصورة مستقلة.

لا بد من خطو

خطوات ثابتة

ومدروسة لضمان

استيفاء المعطيات

المجموعة بعض

المعايير. تتمثل هذه

الخطوات في تمتع

معطيات الفرز

السريع، بحد ذاتها،

"بالقوة".

على معطيات الفرز السريع أن تنسم أيضاً بالصلاحية. يقصد بالصلاحية مدى ملاءمة أي مؤشر مستعمل المفهوم الذي يجري قياسه. يعتبر القياس صالحاً إذا تطابق المؤشر المستعمل في القياس، بدقة وشمولية، مع إطار الغرض موضع القياس ومحتواه. يمكن تفصيل المثل السابق من أجل توضيح هذه النقطة:

يطلب من ثلاثة مراقبين إضافيين (د-ه-و) تحديد حجم الشخص ذاته، أي (ي). قد يبلغ "د" و "هـ" بأن "ي" الذي يبلغ طوله ستة أقدام هو طويل القامة، في حين أن "و" قد يعتبره متوسط الحجم. تكمن المشكلة في التباس مفهوم الحجم واختلاف إمكانات تأويله، أما للبعض الآخر، فقد يتعدى الأمر مجرد الطول، وبالتالي يفترق الحجم إلى الصلاحية. قد يعتبر "د" بأن "ي" ضخم بما أن "ي" أطول بكثير من "د". قد يعتبر "هـ"

أن "ي" هو شخص ضخم بسبب ثقل وزنه الذي يتعدى وزن "هـ". أما "و"، فقد يقرّر أن "ي" متوسط الحجم، كون حجمه وطوله يماثلان تقريباً حجم وطول "هـ" الذي يعتبر نفسه شخصاً متوسط القامة. وبالفعل، يُعدّ مفهوم الحجم مشكلة تهذد الموثوقية والصلاحية.

لهذه الأسباب عيناها، يجب معالجة الاستطلاعات التي تلي خروج الناخبين من مراكز الاقتراع واستطلاعات الرأي الأخرى بحذر شديد. فغالباً ما ينتج عن هذين النوعين من الاستطلاعات تقديرات غير موثوق بها، بخصوص النتائج الفعلية للتصويت في اليوم الانتخابي. يعود سبب الاختلاف إلى أن الاستطلاعات التي تلي الخروج من مراكز الاقتراع تقيس عادة ما يتذكره المقترعون، في حين تقيس استطلاعات الرأي نية المواطن في التصويت ليس إلا. ويبيدي الناس استعداداً لإعطاء تقرير مشوّه حول كيفية تصويتهم وكيف ينوون التصويت، لأسباب يمكن فهمها. أمّا عمليات الفرز السريع فيمكن الوثوق بها، وهي صالحة، بما أن المراقبين يجمعون نتائج الفرز السريع الرسمية من مراكز اقتراع فردية. تقيس عمليات الفرز السريع السلوك، لا ما يتذكره الناخب وما ينوي القيام به. فهي تقيس كيفية انتخاب المقترح بالفعل، لا كيف كان سيبلغ عن صوته إلى أشخاص يجهلهم تماماً.²

أسئلة متكررة

كيف يضمن رؤساء الفرز السريع أن معطيات هذا الفرز موثوق بها وصالحة؟

تبدأ الخطوة الأولى بضمان فهم كلّ الرؤساء والموظفين هذه المفاهيم. يجب أن يعي كل شخص معنى الأمر بأن نتائج الفرز السريع ستقتد من أهميتها في حال عدم الوثوق بالمعطيات (التي يمكن التحقق منها بشكل مستقل) وعدم صلاحيتها (لدى قياس النية). أمّا الخطوة الثانية فتصف العملية بحد ذاتها. فيجب مثلاً اختبار الأسئلة المطروحة في استمارات المراقب للتأكد من صلاحيتها، كما يجب أن تتضمن استمارات المراقب فئات الأجوبة التي تسمح بتبليغ المعطيات على نحو يمكن الوثوق به. أمّا البرامج التدريبية فهي أساسية أيضاً. يجب تصميمها بطريقة تتيح التأكد بأن كلّ المتطوعين يستوعبون المفاهيم التي تشملها المراقبة، وبأن المراقبين كافة سيقضون الحدث ذاته باستعمال الاستمارة ذاتها، وأتباع الطريقة نفسها.³

العينة

ترتبط صلابة معطيات الفرز السريع بكيفية أخذ العينة. فتحدّد العينة الأصوات التي تستعمل كأساس ترتكز عليه نتائج الانتخابات. تظهر الفكرة الأساسية من العينة طرائق مختلفة في الحياة اليومية. مثلاً، يأخذ الكيميائيون بصورة روتينية "عينة" من مركب، ويقومون بتحليلها لإجراء تقارير دقيقة حول خصائص المركب الكيميائية بأكملها. يأخذ الأطباء عينات الدم من مرضاهم عادةً، لتحديد إذا كان تركيب دمهم يسبب لهم المرض. لكن، لحسن الحظ لا يضطر الطبيب إلى سحب كلّ الدم من جسم المريض لمعرفة ما يحتويه بشكل دقيق. فمقاربة الموضوع بهذا الشكل غير عملية

² تقيس عمليات الفرز السريع أيضاً الأوجه النوعية لعمليات التصويت وفرز الأصوات. وكما ذكرنا في الفصل السادس، العنصر النوعي في الفرز السريع، يجب إيلاء اهتمام هائل بطرح الأسئلة لقياس المؤشرات النوعية.

³ تتم معالجة هذه المشاكل بتفصيل أكبر خلال هذا الدليل. وبشكل خاص، يناقش الفصل الرابع، إنشاء شبكة المتطوعين، تصميم استمارات وكتيبات المراقب واستقطاب المتطوعين وتدريبهم. أما الفصل السادس، العنصر النوعي في الفرز السريع، فيقدّم المزيد من التوصيات لتصميم استمارات مراقب الفرز السريع.

وغير ضرورية، بما أن مجرد أخذ عينة واحدة من الدم تكشف كل ما يحتاج الطبيب إلى معرفته حول ما يحتوي دم المريض كله.

تعتمد عينات الفرز السريع على المبادئ ذاتها تماماً. قد تطلب مجموعة مراقبة من المتطوعين مراقبة كل مركز اقتراعي في البلاد، والإبلاغ عن كل نتيجة يحصلون عليها. غير أن هذه الاستراتيجية تتطلب قدراً كبيراً من الموارد، علماً بأن هذه الطريقة غير ضرورية. على غرار عالم الكيمياء والطبيب، تستطيع مجموعات المراقبة معرفة كل ما تحتاج إليه بشأن الناخبين كافة، عبر استعمال عينة صممت بشأن. فهذه الطريقة أسرع وأوفر ثمناً وأكثر عملية.

توفر عينات الفرز السريع أساساً موثوقاً به للوصول إلى تقديرات دقيقة عن السكان كافة.

توفر عينات الفرز السريع أساساً موثوقاً به، للوصول إلى تقديرات دقيقة عن السكان كافة، لأن العينة تمثل مجموعة فرعية من كامل السكان وهي مجموعة ثانوية تظهر مميزاتهم. بالرغم من ذلك، يفترض تصميم العينات أيضاً اتخاذ القرارات التي تملك تأثيراً عميقاً على دقة المعطيات، وأنواع تحليل البيانات الممكنة في الوقت نفسه.

السكان

من الناحية التقنية، يُقصد بالسكان كل الحالات الفردية ضمن إطار حدود ما. غالباً ما لا يُعنى علماء الإحصاءات بتعداد الأفراد. لا تهتم عمليات الفرز السريع بكل فرد يعيش داخل حدود بلد ما. بل تهتم هذه العمليات بالسكان المعنيين وحسب، أي بكل فرد مؤهل للتصويت.

يستثنى من السكان الذين تشملهم عملية الفرز السريع كل الأشخاص الذين ليسوا مؤهلين للاقتراع، لأي سبب قانوني. تملك القوانين الانتخابية في معظم البلدان قواعد واضحة تتعلق بالعمر مثلاً. فالأشخاص الصغيرو السن ليسوا مؤهلين مثلاً للتصويت، بالرغم من أن العمر المحدد للاقتراع يتفاوت بين بلد وآخر. ولكن معظم البلدان تتشابه بامتلاكها متطلبات مواطنة، تتيح للمواطنين الانتخاب في الانتخابات الوطنية فقط.⁴

تبدأ عمليات الفرز السريع بافتراض أن معطيات فرز

الانتقال من العينة إلى السكان

تبدأ عمليات الفرز السريع بافتراض أن معطيات فرز الأصوات موثوق بها وصالحة بحد ذاتها. بمعنى آخر، تفترض عمليات الفرز السريع أن فرز الأصوات الرسمي في مراكز الاقتراع تشكل معلومات صلبة— أي المعطيات التي يجمعها المراقبون من كل مركز لأخذ العينات—. وبالفعل، تستطيع مجموعات المراقبة التحقق من هذا الافتراض بإجراء مراقبة نوعية، ونظامية، لعمليات التصويت والفرز في مراكز الاقتراع.⁵

الأصوات موثوق بها وصالحة بحد ذاتها.

في حال أثبتت المراقبة النوعية لتدابير اليوم الانتخابي أن معطيات الفرز السريع موثوق بها وصالحة، وإذا تم اتباع المبادئ الإحصائية الأساسية،

⁴ من الجدير بالملاحظة أنه يمكن تجاهل طبيعة الانتخابات الديمقراطية حول أهلية الاقتراع من خلال استثناءات غير قانونية وتمييزية، و/أو عبر تلاعب في لوائح الانتخابات الرسمية. لا تعالج عمليات الفرز السريع المشاكل المماثلة غير أنه يجب أن نتطرق إليها نشاطات أخرى تُعنى بمراقبة الانتخابات. راجع مثلاً كتاب: بناء الثقة في عملية تسجيل الناخب، دليل المعهد الديمقراطي الوطني للأحزاب السياسية والمنظمات المدنية (2001).

⁵ يفصل الفصل السادس، العنصر النوعي في الفرز السريع، الإجراءات المثبتة لتقييم نوعية تدابير التصويت والفرز بصورة نظامية.

يمكن عندها الحصول على تقديرات دقيقة لتوزيع التصويت على البلد برمته، على أساس عينة استخرجت بحسب الأصول. من الممكن إجراء تقديرات دقيقة للغاية حول سلوك السكان (كيف ينتخب السكان) على أساس عينة (من النتائج الصادرة عن مراكز الاقتراع المنتقاة) بسبب نظرية الأرجحية.

الأرجحية: قانون الأعداد الكبرى ونظرية الحد المركزي

تتعلق الأرجحية بفرصة حصول حدث ما أو عدم حصوله. من الممكن تقدير أرجحية أحداث مستقبلية مجهولة- كأن تريح البرازيل كأس العالم، أو أن تهطل الأمطار اليوم. صحيح أنه ليس في مقدور أحد أن يعلم مسبقاً ماذا سيحصل، لكن يمكن القيام بتخمين جيد يرتكز على أداء الفريق في أحداث أخرى، أو على الأحوال الجوية في الخارج. من الممكن أيضاً التنبؤ بالأرجحية ارتكازاً على إمكانية حصول شيء. أدرس المثال الإحصائي التقليدي لتخمين الوجه الذي ستسقط عليه العملة النقدية المعدنية التي ترمى في الهواء، إنما بدون خداع:

ثُرِمى العملة في الهواء 100 مرة. عند استعمال القطعة النقدية الصحيحة، ستكون فرص ظهور وجهها (طرة) 50 مرة وظهريها (نقشة) 50 مرة أو ما يعادل هذه النسبة. فلنفترض أن القاعدة نفسها قد اختبرت برمي القطعة النقدية نفسها بعض المرات لا غير. فرمي هذه القطعة 12 مرة في الهواء قد يحدث نتيجة غير متعادلة بالضرورة. فقد تقع العملة على وجهها 9 مرات و3 مرات على ظهرها. وبالفعل، من الممكن في ظروف استثنائية، خلال اثنتي عشرة رمية، أن يظهر وجه القطعة النقدية في كل مرة. ففي الواقع، يمكن أن يحسب المرء، بدقة عالية جداً، أرجحية الحصول على نتيجة غير اعتيادية. فتكون أرجحية الحصول على 12 وجه بصورة متتالية $(1/2)^{12}$ أو واحد من 4096 أو 0.024 بالمئة؛ أي أن فرصة الحصول على اثني عشر وجه (أو ظهر) بشكل متتالي هي واحد من أصل أربعة آلاف وستة وتسعين. تشير نظرية الأرجحية إلى أن التوزع بين الوجوه والأظهر سيتعادل في الأمد الطويل.

يتمثل أحد أوجه نظرية الأرجحية، في العمل، وفقاً لمثل رمي العملة النقدية أعلاه، بنظرية الأعداد الكبرى. تعتبر هذه النظرية بأنه، كلما رُميت العملة في الهواء، كلما زادت الأرجحية بأن يطابق التوزيع الإجمالي للنتائج النهائية (المراقبات) نمطاً متوقعاً ومعروفاً تماماً. أما المعنى المتضمن فواضح: كلما زادت المعطيات في حوزتنا كلما أصبحنا أكيدين من القدرة على تخمين النتائج بشكل دقيق.

يرتكز قانون الأعداد الكبرى الإحصائي على الرياضيات، أما الأمثلة غير الفنية التي يمكن استخلاصها، فهي تمتع الأرقام بالأمان. يظهر مثل ثان نقطة مهمة لفهم أساس منهجية الفرز السريع.

أدرس مجموعة تتألف من 500 طالب يتابعون الدروس ذاتها في الجامعة. يحصل معظم الطلاب على علامة "جيد جداً" و"جيد"، بالرغم من أن بعض الطلاب ينالون علامة "ممتاز"، والبعض الآخر "وسط" أو حتى "ضعيف". لا شك في أن توزيع العلامات نفسه قد لا يتكرر تماماً في الصف

نفسه الذي يحتوي على 10 طلاب أو أقل. والأهم هو أن العلامات الممتازة أو المتدنية بصورة استثنائية ستتمتع بأثر مختلف على معدل العلامات في كامل الصف. ففي الصفوف الصغيرة، يكون لهذه العلامات المتطرفة أثر كبير على التوزيع الإجمالي وعلى معدل الصف، وتؤدي بالتالي إلى انحراف منحني العلامات. لكن أثر أية علامة استثنائية، في الصفوف الأكثر عدداً، سيكون أقل على معدل علامات كامل الصف.

أما المعنى العملي لمثل توزيع العلامات فهو بسيط: كلما زادت كمية المعطيات (عدد مراكز المراقبة)، كلما تقلص أثر أية نقطة معطيات فردية على النتيجة الإجمالية.

يُعرف المبدأ الإحصائي الثاني، الأساسي بالنسبة لمنهجية الفرز السريع، بنظرية الحد المركزي. تعتبر هذه الحقيقة البديهية أنه مع زيادة عدد المراقبات (مراكز العينات) تزيد أرجحية مطابقة توزيع مراكز المعطيات نمطاً معروفاً. فتوزع علامات الصف الذي يتألف من 500 طالب، في علم الفيزياء في البرازيل، بطريقة مماثلة للصف الذي يتألف من 500 طالب أدب في فرنسا، رغم اختلاف العلامات نفسها. في كلتا الحالتين، تتقارب معظم مراكز المعطيات من معدل العلامات.

تعمل هاتان الحقيقتان الإحصائيتان بشكل مشترك - أي قانون الأعداد الكبرى ونظرية الحد المركزي. وهما تشيران معاً إلى أنه:

1. كلما زاد عدد المراقبات (مراكز أخذ العينات) كلما تقلصت أرجحية تأثير علامة فردية استثنائية على المعدل العام؛ و
2. كلما زاد عدد المراقبات، كلما زادت أرجحية توزيع مجموعة المعطيات بمجملها، وفقاً لمنحني عادي (نظرية الحد المركزي).

ينشأ مبدأ عام من هذه القواعد الإحصائية، يتمتع بتضمينات قوية في عمليات الفرز السريع: فكلما زاد عدد المراقبات، كلما زادت إمكانية إجراء تخمينات إحصائية موثوق بها، في ما يتعلق بميزات السكان. غير أن فهم هذا المبدأ يشكل مسألة حاسمة؛ فمن أجل مراعاة هذين المبدأين الإحصائيين، يجب انتقاء حالات العينة بطريقة عشوائية.

العشوائية

يمكن اعتبار العينة، لا كمجرد مجموعة فرعية من السكان وحسب، بل كنسخة مطابقة ومصغرة عن السكان الذين أخذت منهم العينة. يُمكن الاعتبار بأن سكان كل بلد يتمتعون بصفات فردية في بعض المجالات. فيستحيل حصول تشابه تام بين بلدين في توزع بعض الصفات، كاللغة والدين والجنس والمهنة ومستوى التعليم، على السكان. ولا شك في أن امتلاك فرد ما سيارة، أو السكن في المدينة بدل البلدة، أو الحصول على فرصة عمل، أو تربية كلب أليف، يساهم في سمة الخبرة الشخصية الفريدة. من المستحيل وضع لائحة نهائية وشاملة تتضمن كل خاصية

كلما زاد عدد المراقبات، كلما زادت إمكانية إجراء تخمينات إحصائية موثوق بها في ما يتعلق بميزات السكان.

تميّزنا كأفراد، ناهيك عن كامل السكان. فعلاقات العناصر في ما بينها، أي التي ينبغي تسجيلها في وثائق، كثيرة جداً. لكن لحسن الحظ لا تتطلب عمليات الفرز السريع هذا الأمر، كما أنها لا تهتم بكل ما يميّز الناس عن بعضهم البعض، بل تهتم حصرياً بالعناصر التي تؤثر، بشكل يمكن إثباته، على توزيع الأصوات ضمن السكان المؤهلين للاقتراع.

يجب انتقاء النقاط
العينية من السكان،
على نحو عشوائي،
وبهذه الطريقة فقط
كي تمثل العينات
النهائية كامل السكان.

يجب انتقاء النقاط العينية من السكان، على نحو عشوائي، وبهذه الطريقة فقط كي تمثل العينات النهائية كامل السكان. تعني العشوائية عملياً التماثل التام بين أرجحية اختيار أية نقطة فردية من العينات التي تؤخذ من السكان، وأرجحية اختيار أية نقطة أخرى يتم اختيارها. وللأسباب التي تطرقنا إليها سابقاً، يشير قانون الأعداد الكبرى ونظرية الحد المركزي إلى أنه، كلما كبر حجم العينة المستخرجة، كلما مثلت هذه العينة مواصفات السكان إلى حدٍ أكبر.

التجانس والتغاير

لا تتطلب العينات الموثوق بها كميات هائلة من المعلومات المفصلة حول خصائص كامل السكان الاجتماعية. غير أنه لا بدّ من معرفة إذا كان السكان المعنّون متنوعين (متغايرين) أم لا (متجانسين). تخلف تقييمات التغاير والتجانس آثاراً شديدة على كيفية أخذ عينات السكان بطريقة يمكن الوثوق بها.

أما طرائق التحقق من مستوى التغاير أو التنوع في أية مجموعة سكانية، فعديدة. وتجدر الإشارة إلى تأثير التركيبة الإثنية والدين واللغات على التغاير. غير أنّ المشكلة الأولى التي تواجه عمليات الفرز السريع لا تنحصر فقط بمستوى التغاير الإثني والسياسي في السكان. إذ تكمن المسألة الأساسية في عمليات الفرز السريع في تأثير التغاير بشكل فعال على سلوك التصويت. في حال فضل 80 % بالمئة من السكان أحد المرشحين، فإنّ ذلك يعني أنّ السكان متجانسون إلى حدٍ ما، بغضّ النظر عن تنوعهم الديني واللغوي والإثني. ونسجاً على المنوال نفسه، يُعدّ السكان متغايرين نسبياً في حال انتهاء المعركة الانتخابية، وتعادل الأصوات بين مرشحين أو أكثر.

يقضي الارتفاع في
تغاير النسبة المقترعة
من السكان زيادةً في
العينة من أجل

أما سوء التقدير الشائع فهو اعتبار السكان الذين يُسمون بالتنوع الاجتماعي غير متجانسين في التصويت. لكنّ مجرد تغاير السكان الاجتماعي لا يفرض تغايرهم في الانتخابات. فعلى سبيل المثال، تتمتع الهند بتعددية اللغات والأديان، إنّما هي متجانسة نسبياً في حال أردنا إنشاء عينة من السكان المقترعين.

الحصول على تقدير
دقيق لسلوك
التصويت.

يقضي الارتفاع في تغاير النسبة المقترعة من السكان زيادةً في العينة، من أجل الحصول على تقدير دقيق لسلوك التصويت. من شأن هذه المسألة أن تتوضّح جيّداً، بفضل المقارنة بين أحجام العينات اللازمة والمأخوذة من ثلاثة بلدان ذات أحجام سكانية مختلفة، هي كندا والولايات المتحدة وسويسرا.

الصورة 1-5:

أحجام العينات المطلوبة
للولايات المتحدة وكندا
وسويسرا

السكان	الولايات المتحدة	كندا	سويسرا
263,814,032	28,434,545	7,084,984	
هامش الخطأ	%2-/+	%2-/+	%2-/+
حجم العينة المطلوبة	1200	2400	4300

كما تظهر الصورة 1-5، لا تُحدّد خصائص السكان الإثنية التغيرات. يتحدّد التغيرات عبر إمكانية فوز أحد المرشحين بغالبية الدعم الانتخابي. ففي نظام ثنائي الأحزاب كما في الولايات المتحدة، يسهل غالباً متابعة المعركة الانتخابية كما يمكن التنبؤ بالنتيجة بشكل أسهل- فعلى المقترعين الانتقال بين خيارين ليس إلا. أما في سويسرا، فيزيد عدد الأحزاب الأكبر من حدة المنافسة. تحظى الأحزاب السياسية السويسرية بدعم مجموعات متنوعة اللغات والأديان. لا بل إنّ كندا نفسها، بأحزابها الرسمية الخمسة، تعدّ أقلّ تغيّراً من سويسرا.

أما الاستراتيجية
الأكثر أماناً فهي التي

تقوم بالافتراض

التقليدي أنّ النسبة

المقترعة من السكان

هي متغيرة.

تبيّن الصورة 1-5 أيضاً مبدأ يمتّ بصلّة إلى هذا الموضوع. يتحدّد حجم العينة المطلوبة وفق مستوى التجانس المتوقع في نتائج التصويت، لا وفق عدد السكان الإجمالي في البلد. تقتضي هذه البلدان ذات الأعداد السكانية المختلفة عينات مختلفة الأحجام، من أجل الحفاظ على هامش خطأ يتراوح بين (+/-2%). وبالفعل يبدو أنّ البلد ذا العديّة السكانية الأكبر يتطلب العينة الأصغر. وفي الواقع تُنسب الاختلافات في حجم العينة المطلوبة إلى الاختلافات في تجانس المجموعات السكانية الثلاث.

ومن العسير في الواقع إيجاد المعلومات التي يمكن الوثوق بها حول تغيّرات النسبة المقترعة من السكان، أو تجانسها، في العديد من البلدان. أما الاستراتيجية الأكثر أماناً في الظروف المماثلة، فهي تلك التي تتطلب تخمينات، فنفترض الافتراض التقليدي أنّ النسبة المقترعة من السكان هي متغيرة. وكما سيُضح لاحقاً، يخلف هذا الافتراض أثراً عميقاً على كيفية احتساب عينة من عملية الفرز السريع.

كلّما استوجبت الثقة

أن يعكس توزيع

العينة التوزيع

السكاني، كلّما كبر

حجم هذه العينة.

مستويات الثقة: تحديد العلاقة بين العينة والسكان

لعلّ معلومة إضافية واحدة تخلف أعرق الأثر على كيفية تقدير علماء الإحصاءات السكان على أساس العينة، وهي مستوى الثقة. تعني مستويات الثقة طريقة مقارنة معطيات العينة مع السكان. وكلّما استوجبت الثقة أن يعكس توزيع العينة التوزيع السكاني، كلّما كبر حجم هذه العينة. ويعود هذا الأمر إلى أنّ تأثير النتائج الفردية الاستثنائية على التوزيع سيقف في العينات الكبيرة.

وبما أن عواقب نتائج
الفرز السريع غير
الدقيقة قد تكون
وخيمة، تقضي
الإجراءات المعيارية
في مراقبة الانتخابات
بأخذ عينات بمستوى
ثقة يبلغ 99 بالمئة.

اعتاد الإحصائيون الاعتماد على مستوى ثقة بنسبة 92 بالمئة. يعبر مستوى الثقة، تقنياً ومثوياً، عن الأرجحية التي تؤكد على أن معدل العينة سيوفر تقديراً دقيقاً عن معدل السكان. وبالتالي، تشير نسبة 95 بالمئة من مستوى الثقة إلى أن 95 بالمئة من كل معدلات العينات ستطابق معدل السكان بالفعل. وبما أن عواقب نتائج الفرز السريع غير الدقيقة قد تكون وخيمة، تقضي الإجراءات المعيارية في مراقبة الانتخابات بأخذ عينات ذات صفات تقليدية، أي بمستوى ثقة يبلغ 99 بالمئة.

أخذ العينات

تشمل الطريقة العملية لأخذ عينة للفرز السريع مزيجاً من القرارات، فهي تتضمن:

- التعرف على وحدة التحليل؛
- تحديد هامش الخطأ ومستويات الثقة؛
- تحديد نوع العينة العشوائية الأكثر ملاءمة؛ و
- تقدير عوامل التصحيح لنسب استخراج العينات والامتناع عن التصويت.

وحدة التحليل

تشير وحدة التحليل إلى المادة محور الدراسة. فإذا كان الهدف تعميم الدراسة على كامل السكان، يمثل الفرد غالباً وحدة التحليل. غير أنه من الممكن في بعض الحالات تعميم العينة على السكان، باعتماد مجموعة أكبر كوحدة تحليل، كالأسرة أو مجمع سكني في المدينة.

تستعمل عمليات الفرز

في عمليات الفرز السريع، يكون الهدف تقدير توزيع أصوات المواطنين بين الأحزاب السياسية. غير أن صوت الفرد في الانتخابات الديمقراطية يكون سرياً، وبالتالي لا يمكن أن يمثل وحدة التحليل. بدلاً من ذلك، تستعمل عمليات الفرز السريع، نموذجياً، النتائج الرسمية في مركز اقتراعي فردي كوحدة التحليل. يعود هذا الأمر إلى أن المركز الاقتراعي هو أصغر وحدة تحليل تُجمع فيه الأصوات الفردية، وإلى أن القواعد الانتخابية تقترض عادة أن يجري الفرز الرسمي في المركز الاقتراعي.

هامش الخطأ: إلى أي حد علينا أن نكون دقيقين؟

يشكل هامش الخطأ أحد أهم المعلومات التي تؤخذ في الحسبان لدى إعداد العينة. يشير هامش الخطأ، الذي يعبر عنه كنسبة مئوية، إلى سلسلة القيم المحتملة في أية مراقبة. يسلط المثل التالي الضوء على المفهوم:

تشير النتائج الصادرة عن مركز اقتراعي واحد بأن 48 بالمئة من الأصوات تدعم المرشح "أ". فإذا كان هامش الخطأ المئبي بنسبة خمسة بالمئة، فيدعونا سبب وجيه إلى الوثوق بأن النتائج الفعلية التي حصل عليها المرشح "أ" تتراوح ما بين 43 و 53 بالمئة لدى أخذ كافة المقترعين بعين الاعتبار من مجموع السكان.

تصمّم المنظمات المدنية التي تجري عمليات الفرز السريع عينات الفرز السريع، نموذجياً، للحصول على هامش خطأ زائد أو ناقص 0.5 بالمئة (+/-0.5%). قد يظهر سبب عرضي (كتوقع تقارب عدد الأصوات) يحثّ على اختيار هامش للخطأ أكثر صرامة. فهامش الخطأ المطلوب رهن بدرجة الدقة التي تستوجبها التقديرات.

تصمّم المنظمات المدنية التي تجري عمليات الفرز السريع عينات الفرز السريع، نموذجياً، للحصول على هامش خطأ زائد أو ناقص 0.5 بالمئة (+/-0.5%).

يُقاس هامش الخطأ باستعمال المعادلة التالية:

$$ME = \frac{s}{\sqrt{n}} * z$$

حيث تمثّل

ME = هامش الخطأ

s = الانحراف الطبيعي (افتراض 0.5)

n = حجم العينة

z = القيمة z لمستوى الثقة المنتقاة (تكون 1.96 لنسبة 95%، و 2.58 لنسبة 99%)

تحتوي أية مجموعة من المعطيات، وهي مجموعة مراقبات تستند إلى مركز من العينات، على خاصيتين على الأقل. تتمتع هذه المعطيات بميل مركزيّ تتجمع حوله معظم النتائج، وهي تحتوي أيضاً على تباين أو تشتت. يقصد بالتباين مدى توزيع المراقبات، بشكل موسّع أو محدود. أما طرائق قياس الميل المركزي والتشتت، فعديدة وهي تتصل بحسابات هامش الخطأ.

قياس الميل المركزي

إنّ طريقة قياس الميل المركزي الأكثر شيوعاً هي المعدل، أي معدل قيمة المراقبات المسجّلة كافة. يُشتقّ المعدل من خلال إضافة قيم كل مراقبة إلى مجموعة المعطيات، ثم اقتسامها على عدد المراقبات. يوضح المثل التالي هذه العملية:

يبلغ معدل الأرقام التالية: 1، 3، 4، 6، 7، 9 العدد 5 الذي نحصل عليه بالشكل التالي: $30 = 9+7+6+4+3+1$ ، وبما أنّ عدد المراقبات هو 6، تكون المعادلة $5 = 30 \div 6$.

تتعدّد الأساليب لقياس الميل المركزي لأي نوع من المعطيات. يشير النمط مثلاً إلى العدد الأكثر تكراراً في أية مجموعة من المعطيات. ففي المجموعة التالية من الأعداد: 1، 3، 3، 3، 5، 6، 7، تتكرّر ملاحظة العدد 3. لكن لا بدّ من الملاحظة بأنّ معدل المجموعة عينها من الأعداد هي 4 $[4 = 7 \div (7+6+5+3+3+3+1)]$.

أما القياس الثالث للميل المركزي فهو العدد المتوسط. يظهر هذا العدد في وسط مجموعة معينة من المراقبات. ففي المجموعة التالية من المعطيات: 1، 3، 6، 7، 8،

8 و10، يعتبر العدد الذي يظهر وسط المراقبات هو 7؛ وتجدر ملاحظة ثلاثة أعداد أصغر من 7 وثلاثة أخرى أكبر من 7. أما نمط هذه المجموعة فهو 8 لأن العدد 8 هو الذي يتكرر. أما معدل المجموعة عينها فهو 6.14. يبلغ الإحصائيون عادة عن المعدل بدلاً من العدد المتوسط أو النمط كونه القياس الأكثر فائدة للميل المركزي.

قياس التشبث

تحتوي المعطيات على ميزة ثانية هي قياس التشبث، لمعرفة كيفية توزيع قيم المراقبة، أي مدى توسعها أو محدوديتها. في ما يخص المثل أعلاه، من الواضح أن أية مجموعة من المعطيات ستملك معدلاً خاصاً بها. غير أن هذا المعدل لا يعطينا معلومات حول كيفية تشبث القيم التي تمت مراقبتها أي مدى توسعها أو محدوديتها. تتشاطر المجموعات التالية المعدل الحسابي نفسه، وهو 3:

2، 2، 3، 4، 4

99-، 99، 3، 99، 99

إن طريقة توزيع المجموعتين السابقتين من المعطيات مختلفة تماماً. تكمن إحدى الطرائق المثبتة لإظهار الاختلاف بين المجموعتين، في معرفة تسلسل الأعداد. فأصغر عدد في المجموعة الأولى هو 2 أما العدد الأكبر فهو 4. أما التسلسل الذي نحصل عليه فهو 4 ناقص 2 يساوي 2. في المجموعة الثانية، العدد الأصغر هو 99 سلبى والعدد الأكبر هو 99 إيجابى، فيكون التسلسل عندها 99 إيجابى ناقص 99 سلبى، أو 198.

من الواضح أن تسلسل المجموعتين المختلفتين يبرز أحد أوجه الاختلافات الأساسية بين مجموعتي الأعداد. وبالرغم من ذلك، لا يُعنى التسلسل إلا بعددين، الأكبر والأصغر، متجاهلاً نقاط المعطيات الأخرى كافة. كما يمكن إيضاح المزيد من المعلومات حول تبدد المراقبات ضمن مجموعة المعطيات من خلال نوع آخر من القياس، هو التباين.

يشير التباين، في مفهومه غير التقني، إلى معدل المسافات كافة بين كل قيمة مراقبة ومعدل قيم المراقبة كافة. يأخذ التباين في الحسبان معدل مجموعة من المعطيات، وعدد المراقبات، بالإضافة إلى كل مراكز المعطيات بحد ذاتها. فيشمل، نتيجة لذلك، كل المعلومات اللازمة لشرح انتشار مجموعة المعطيات. يمكن تحديد أية مجموعة من المراقبات في أربع خطوات:

1. احتساب معدل مجموعة المعطيات.
2. احتساب المسافة بين كل مركز معطيات وبين المعدل، وتربيع المسافة.
3. إضافة المسافات المربّعة كلها معاً.
4. اقتسامها على عدد المراقبات.

فتكون المعادلة كالتالي:

وفقاً لمجموعة من المعطيات تحتوي على المراقبات $X_1, X_2, X_3, \dots, X_n$

$$s^2 = \frac{(x_1 - x)^2 + (x_2 - x)^2 + (x_3 - x)^2 \dots (x_n - x)^2}{n-1}$$

حيث تمثّل

$s^2 =$ التباين

$X_1, X_2, X_3, \dots, X_n$ هي المراقبات

x هو المعدّل

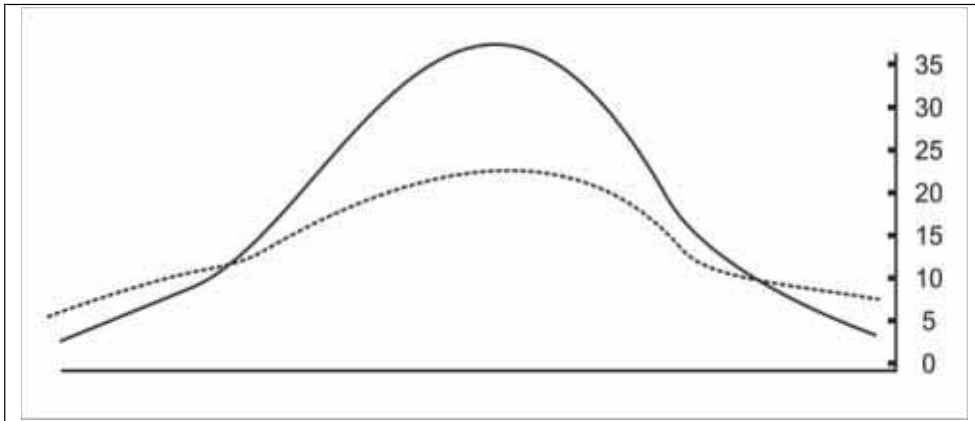
n هو عدد المراقبات

$$s^2 = \frac{\sum (x_n - x)^2}{n-1}$$

وبصيغة مختصرة، تصبح المعادلة:

الانحراف القياسي هو الجذر المربع للتباين. يعتمد الإحصائيون عادة على الانحراف القياسي لأنه يظهر التباين في وحدات قياسية يمكن مقارنتها بشكل ملحوظ. كلما زاد الانحراف القياسي بالنسبة لأية مجموعة من المعطيات، كلما ابتعدت المعطيات عن المعدّل. وكلما تقلص الانحراف القياسي، كلما اقتربت مراكز المعطيات الفردية من المعدّل بشكل وثيق.

يجب أخذ مفهوم إضافي للقياس بعين الإعتبار: مفهوم التوزيع الطبيعي. تظهر المناقشة السابقة بأنّه، في كلّ مجموعة من المعطيات، تتجمّع مراكز فردية من المعطيات حول معدّل أو نقطة متوسطة. أما الطريقة الأخرى للتعبير عن الفكرة نفسها فتتصّل على معرفة أية نسبة من المراقبات تقع ضمن الانحراف القياسي للمعدّل. إذا كانت مجموعات المعطيات كبيرة بما فيه الكفاية، ومتطابقة مع مبادئ العشوائية، سيتطابق تشتت قيم المعطيات مع ما يسمى بالتوزيع الطبيعي. يملك التوزيع الطبيعي بعض الخصائص المعروفة: المنحنى الطبيعي كما يظهر في الصورة 2-5، فهو جرسيّ الشكل ومتماثل الجانبين، ويتلاقى فيه المعدّل والنمط والعدد المتوسط.



الصورة 2-5:
منحنيا التوزيع الطبيعي

يحدّد حجم التباين شكل التوزيع الفعلي الدقيق. ففي تحقيق سبيل أهداف الفرز السريع، من الضروري أن تتمتع كلّ مجموعة من المعطيات التي تتطابق مع منحنى التوزيع الطبيعي بالخصائص القياسية عينها. وهي: تدرج 68.3 بالمئة من كلّ القيم التي تمت مراقبتها ضمن انحراف قياسي من المعدل، و95.4 بالمئة من إجمالي النتائج ضمن انحرافين قياسييين من المعدل، و99.7 من مجمل النتائج ضمن ثلاثة انحرافات قياسية من المعدل. لن تتطابق مجموعات المعطيات كلّها مع هذا النموذج الدقيق. وفي حال احتوت المعطيات على تباين شديد، سيكون المنحنى منبسّطاً إلى حد ما. أما في حال ظهور بعض الاختلاف، سيصبح المنحنى أكثر تقوّساً.

يمكن الإشارة إلى المسافة الفاصلة عن المعدل التي تعبّر عنها الانحرافات القياسية بالمجموع Z أو القيم الحرجة. تحتوي معظم الكتب الإحصائية التعليمية العادية جدولاً بقيم Z حول التوزيع الطبيعي؛ فلا يضطر المحللون إلى احتساب القيم Z كلّ مرّة يواجهون فيها مجموعة من المعطيات. إذا تضمّنت المعطيات مساحة ثقة بنسبة 95 بالمئة (سيشمل 95 بالمئة من كلّ معدّلات العينات معدل السكان)، يتّضح عندئذٍ أنّ النتائج ستندرج ضمن 1.96 انحراف قياسي عن المعدل. ونسجاً على المنوال نفسه، يشير مستوى الثقة بنسبة 99 بالمئة إلى أنّ 99 بالمئة من النتائج كافة (حيث يشمل معدل العينة معدل السكان) تشكّل 2.58 انحرافات قياسية عن المعدل. في هذه الحالات، تمثل القيم 1.95 و2.58 القيم الحرجة، أو القيم Z ، بالنسبة لمستويات الثقة 95 و99 بالمئة على التوالي.

تؤثر القرارات حول أي

هامش خطأ يمكن

احتماله في الفرز

السريع، وبشكل مباشر،

على الحسابات التي

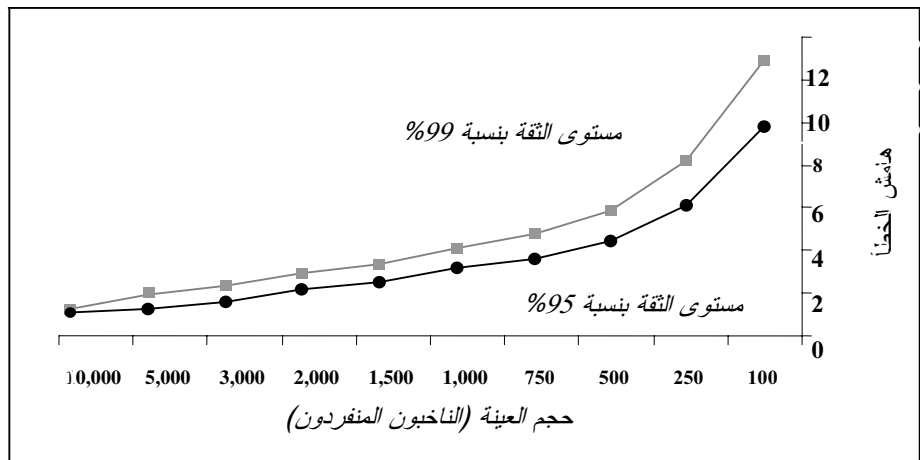
تُجرى لتحديد الحدّ

الأدنى المطلوب من

حجم العينة.

يتطلب احتساب هامش الخطأ الاعتماد على الانحراف القياسي وعلى القيم Z . يشمل الانحراف القياسي والقيم Z ، دورها، قياس الميل المركزي والتشتت ومستويات الثقة. وتفاوت هوامش الخطأ وفقاً لمستويات الثقة وأحجام العينة، كما توضح الصورة 5-3. بصورة عامة، كلما ارتفع مستوى الثقة، كلما زاد هامش الخطأ. وكلما كبر حجم العينة، كلما تقلص هامش الخطأ. تؤثر القرارات بشأن هامش خطأ المحتمل في الفرز السريع تأثيراً مباشراً على الحسابات التي تُجرى لتحديد الحدّ الأدنى من حجم العينة المطلوبة.

الصورة 5-3:
هوامش الخطأ وأحجام
العينات



أنواع العينات

تشمل العينات نوعين أساسيين: عينات الأرجحية وعينات عدم الأرجحية. تتطابق عينات الأرجحية مع مبادئ العشوائية، وهي بالتالي تمثل مجمل السكان. تجدر الإشارة إلى أن عمليات الفرز السريع تستعمل دائماً عينات الأرجحية.

لا تختار عينات عدم الأرجحية مراكز العينات عشوائياً، فيبقى مدى تمثيلها للسكان عامةً مجهولاً. تعدّ عينات عدم الأرجحية مفيدة في بعض الظروف. فهي غير مكلفة ويمكن تشكيلها وإجراؤها بصورة أسهل من عينات الأرجحية. يتم اختيار حالات العينة على أساس سهولة دراستها أو بعثها على الراحة. يقف مثلاً مراسل محطة تلفزيونية خارج ملعب بايسبول، ويسأل المعجبين إذا استمتعوا بالعبة. توفّر هذه الاستراتيجية مشاهد سريعة ومثيرة للاهتمام في عملية البحث، غير أنها لا توفّر معلومات موثوقاً بها حول عدد الحاضرين الإجمالي داخل ملعب الباييسبول.

في ما يتعلق بأهداف الفرز السريع، تقلص عدد عينات عدم الأرجحية لاستحالة الوثوق بها، عند تعميمها على السكان كافة. من هنا، لا تشكل المعطيات التي تؤدي إليها هذه العينات تقديرات يمكن الوثوق بها حول مميزات السكان. ففي حال تمّ تشكيل العينة من مراكز الاقتراع في العاصمة وحسب، لا شك في أن النتائج ستكون مختلفة عن النتائج الصادرة عن المراكز الاقتراعية الموجودة في المناطق الريفية. فالأشخاص الذين يستخرجون المعطيات من الأماكن الملائمة التي يسهل الوصول إليها لا يستعملون المعطيات التي تمثل السكان بمجملهم.

أسئلة متكرّرة

هل يستطيع الإحصائيون تشكيل عينات الفرز السريع التي تجمع بين فوائد عينة الأرجحية وفوائد عينة عدم الأرجحية؟

لا على الإطلاق. تكون العينة إما عينة أرجحية، وفي هذه الحالة تتطابق بالكامل مع مبادئ العشوائية، وإما عينة عدم أرجحية. يؤدي المزج بين عناصر من كلا تقنيّتي العينات إلى تشكيل عينة عدم الأرجحية. ونتيجة لذلك، يجب تفادي الرغبة في استبدال المراكز الاقتراعية "الملائمة" بأخرى يستحيل دخولها. فلا يمكن استبدال مراكز العينات ببعضها البعض، وإلا انتهكت هذه الاستراتيجية مبدأ العشوائية— وهو ينصّ على تمثّل مراكز العينات كافة بأرجحية اختيارها نفسها. بدلاً من ذلك، يترتب على منظمي الفرز السريع أن يؤمّنوا وصول المراقبين الذين خضعوا للتدريب إلى مراكز الاقتراع نفسها المذكورة في العينة، حتى لو كانت تقع في مناطق نائية، بالإضافة إلى الحصول على معلومات من هذه المراكز. أما إن جمعت المعطيات ونقلت من مركز اقتراعي أكثر ملائمة، فقد يعرّض ذلك موثوقية عملية الفرز السريع بكاملها وصلاحيّتها للخطر.

يجب أن تستعمل
عمليات الفرز
السريع، بصورة
دائمة، عينات
الأرجحية من أجل
الحصول على نتائج
تمثّل مجموعة محدّدة
من السكان.

يجب أن تستعمل عمليات الفرز السريع، بصورة دائمة، عينات الأرجحية من أجل الحصول على نتائج تمثل مجموعة محدّدة من السكان. هناك أنواع عدّة من عينات

الأرجحية، ويمكن أن يمثل كلُّ منها السكان بدقة، بالاعتماد على طرائق مختلفة. لعلّ النوعين الأكثر شيوعاً من عينات الأرجحية هما العينة العشوائية العامة والعينة العشوائية الطبقيّة.

لا يمكن أن يبدأ أخذ
العينات في عملية
الفرز السريع إلا بعد
توفر لائحة دقيقة
وشاملة بكلّ المراكز
الاقتراعية.

العينات العشوائية العامّة

وفق نموذج العينة العشوائية العامّة، يجري اختيار وحدات التحليل عشوائياً الواحدة تلو الأخرى من كل السكان. تتبجح هذه الطريقة لكلّ وحدة من السكان فرصة متساوية في الاندراج ضمن العينة. ولكن لا بدّ من وجود لائحة دقيقة بوحدات التحليل كافة، من أجل أن تحظى كلّ وحدة تحليل بفرصة متساوية للاندرراج ضمن العينة.

يشير الإحصائيون إلى لائحة بكلّ أعضاء السكان كإطار لأخذ العينات. في حالة الفرز السريع، تتمثل وحدة التحليل بالمركز الاقتراعي؛ وبالتالي لا يمكن أن يبدأ تشكيل العينات في عملية الفرز السريع، إلا بعد توفر لائحة دقيقة وشاملة بكلّ المراكز الاقتراعية.

العينات العشوائية الطبقيّة

تعتمد العينة العشوائية الطبقيّة مبادئ العشوائية ذاتها التي تطبقها العينة العشوائية العامّة. غير أنّ أطر العينات التي تؤخذ منها مراكز العينات تتألف من طبقات سكانية محدّدة مسبقاً، وحصرية بشكل متبادل. مثال على ذلك:

يرمي مشروع معيّن إلى استعمال عينة من ألف طالب من أجل تعميم حالة عشرين ألف طالب في الجامعة، علماً بأنّ نصفهم لم يحصلوا على الشهادة بعد، فيما يتألف النصف الآخر من الطلاب المتخرجين. وفي حين تختار مقارنة العينة العشوائية العامّة ألف مركز عينيّ من أصل لائحة تضمّ عشرين ألف طالب، تتبجح مقارنة العينة الطبقيّة خطوتين. في الخطوة الأولى، تقسم لائحة كلّ الطلاب إلى مجموعتين (طبقتين)، تشمل المجموعة الأولى الطلاب غير المتخرجين، أمّا المجموعة الثانية فتشمل الطلاب المتخرجين. وفي خطوة تالية، تختار المقارنة 500 حالة من الطبقة الأولى (غير المتخرجين) ومن ثمّ 500 حالة من الطبقة الثانية (المتخرجين).

يعني اللجوء إلى
الطريقة الطبقيّة أنّ
النتيجة النهائية ستنتج
عينة كاملة تعكس تماماً
توزيع الحالات بين
مجمّل السكان.

وفقاً للمقارنة الطبقيّة، ما زال اختيار كلّ حالة يحترم معيار العشوائية: فأرجحية اختيار كلّ حالة في كلّ طبقة تكون نفسها بالضبط (في المثلّ أعلاه 1 من أصل 20). غير أنّ اللجوء إلى الطريقة الطبقيّة يعني أنّ النتيجة النهائية ستنتج عينة كاملة تعكس بشكل دقيق توزيع الحالات بين مجمل السكان. وفي الواقع، تُحدّد عملية وضع الطبقات سلفاً توزيع الحالات ضمن طبقتين أو عدّة طبقات.

يمكن أن تكون الطبقيّة مفيدة وفق طريقة أخرى. لا تملك بعض مجموعات المراقبة الموارد الكافية لإجراء مراقبة على صعيد البلد بكامله. في هذه الحالة، قد تضطر مجموعة المراقبة إلى حصر مراقبتها ضمن طبقة محدّدة من البلد، كالعاصمة أو أية منطقة ساحليّة. حينذاك، يمكن

أن تعمّم مجموعة المراقبة، عبر الاستعانة بمجموعة من مراكز العينات التي اختيرت من الطبقة عشوائياً، نتائج مراقبة على مجمل الطبقة التي تغطيها مجموعة المراقبة.

أسئلة متكرّرة

بما أنه يبدو أن معظم السكان منقسمون إلى طبقات، لماذا لا تُستعمل العينات الطبقيّة في إطار التصويت؟

يبدو للوهلة الأولى أن المقاربة الطبقيّة هي مقاربة مثاليّة، ولكن يحول سببان دون ملامتها الوضع السائد. أولاً، إذا كانت عمليّات الفرز السريع تستعمل الافتراضات التقليديّة بخصوص هوامش الخطأ ومساحات الثقة، من الأرجح أن تكون العيّنة كبيرة. وبسبب نظريّة الأرجحية، من الواضح أن العينات الكبيرة ستخلص إلى إنتاج نسخ متطابقة وتقليديّة عن مجمل السكان، من دون تقسيمهم إلى طبقات حتى. ثانياً، تفترض الطبقيّة إمكانية الثقة بالمعلومات المتعلقة بميل المواطنين إلى التصويت ضمن الطبقات. حول أيّ خط بالضبط يجب رسم معيار الطبقيّة؟ قد لا تكون المعلومات اللازمة للحكم على هذا الأمر موثوقاً بها في العديد من البلدان. ليست الانتخابات السابقة دليلاً يوثق به، لا سيّما إذا كانت نتائجها مثيرة للشكوك. تكمن الاستراتيجية الأفضل في الإشارة إلى الطبقات لاحقاً، من خلال التحقق من توزيع الحالات التي أخذت من عيّنة عشوائيّة عامّة، مقابل طبقات السكان بعد أخذ العيّنة. إذا في حال كان 40 بالمئة من نسبة السكان المقترعين يعيشون في عاصمة البلد، يجب أن تكون 40% من مراكز العينات العشوائيّة جزءاً من العاصمة في نهاية المطاف.

تحديد حجم العيّنة

من أجل تحديد حجم عيّنة عمليّة الفرز السريع (أي عدد مراكز الاقتراع الذي يجب أن تشملها العيّنة)، يتنقل المحللون عبر مراحل متعدّدة. فيحدّدون حجم السكان المعنيين (عدد الناخبين المؤهلين) ومستوى التجانس بين السكان، كما يقومون باختيار مستوى الثقة اللازم وهامش الخطأ. وفي المرحلة التالية، يحسب المحللون حجم العيّنة على النحو التالي:

$$n = \frac{P(1-P)}{\frac{z_{99\%}^2}{N} + \frac{P(1-P)}{N}}$$

حيث تمثّل

$$\begin{aligned} n &= \text{حجم العيّنة (عدد المقترعين المؤهلين للتصويت)} \\ P &= \text{مستوى التجانس المقتر في السكان (بين 0 و 1، أي 50\%)} \\ &= 0.5 \\ \sum &= \text{هامش الخطأ (بين 0 و 1، أي 0.32\% = 0.0032)} \\ z_{99\%} &= \text{مستوى الثقة في حالة التوزيع الطبيعي (99\% في هذه الحالة)} \\ N &= \text{حجم مجمل السكان} \end{aligned}$$

يوضح مثل عملية الفرز السريع التي أجريت أثناء الانتخابات الرئاسية في البيرو، سنة 2001، المراحل المذكورة أعلاه:

كان حجم مجمل السكان المعنّيين (عدد المؤهلين للتصويت) في البيرو 14,570,774. وقد كان من المفترض ألا يكون السكان متجانسين - أي من المتوقع أن تكون المعركة بين كلا المرشحين شديدة أي متقاربة في النقاط، لذا تمّ تحديد مستوى تجانس السكان بنسبة 50% (0.5). تمّ اختيار هامش خطأ بنسبة 0.32 بالمئة ومستوى ثقة بنسبة 99 بالمئة. وبغية إجراء الحسابات، عبّر المسؤولون عن نسبة التجانس بقيمة تتراوح بين 0 و 1، وهو مماثل لهامش الخطأ. ثمّ حدّدوا مستوى التجانس المتوقع بنسبة 50 بالمئة، في سياق افتراض هو الأكثر تحفظاً؛ وأشاروا إليه بالعدد 0.5 ضمن المعادلة، أما هامش الخطأ، ونسبته 0.32 بالمئة (من بين المئة بالمئة الممكنة)، فأشير إليه بالعدد 0.0032. وسرعان ما أدرجت هذه القيم في المعادلة على الشكل الآتي:

$$\begin{aligned}
 N &= \frac{0.50(1 - 0.50)}{\frac{(0.0032)^2}{(2.58)^2} + \frac{0.50(1 - 0.50)}{14,570,774}} \\
 &= \frac{0.25}{\frac{0.000010}{6.6564} + \frac{0.25}{14,570,774}} \\
 &= \frac{0.25}{0.000001515 + 0.000000017} \\
 &= \frac{0.25}{0.000001532} \\
 &= 163,185
 \end{aligned}$$

في هذه المرحلة، يعرف المحللون عدد الناخبين الذين يجب استشارتهم. غير أنّ وحدات التحليل لا تتألف من ناخبين فرديين، بل هي مراكز اقتراعية. بالتالي، تتمثل الخطوة التالية بتحديد عدد المراكز الاقتراعية التي يجب اختيارها من أجل تمثيل عدد الناخبين المطلوب. يمكن متابعة احتساب أرقام حالة البيرو من أجل توضيح هذه النقطة:

شملت انتخابات البيرو، كمعدل وسطي، حوالي 160 ناخباً لكل مركز اقتراع. بالتالي، قُسمت العينة المؤلفة من 163,185 ناخباً مؤهلاً على عدد الناخبين في كلّ مركز (160)، من أجل تحديد عدد المراكز الاقتراعية في العينة الخاصة بنا (1020). وبالتالي، تألفت العينة المستخدمة في عملية الفرز السريع التي أجريت سنة 2001، في البيرو، من 1020 مركز اقتراع.

اختيار مراكز أخذ العينات

ما إن يُعرف حجم العينة العشوائية المطلوبة، حتى يتم اختيارها من إطار العينات الكامل. بهدف إجراء عمليات الفرز السريع، يتم اختيار مراكز الاقتراع (نقاط أخذ العينات) من اللائحة الكاملة بمراكز العينات (إطار العينات). أما الطريقة الأسرع لإجراء هذا الاختيار فتقوم على استعمال برنامج حاسوبي عشوائي. كما يمكن تنفيذ هذه المهمة من دون الاستعانة بالحاسوب. تشمل المرحلة الأولى اقتسام مجمل عدد مراكز الاقتراع على عدد المراكز الاقتراعية المطلوب. أما المرحلة الثانية، فتتطلب تحديد نقطة بداية عشوائية. مرةً أخرى، يمكننا الاستعانة بالأعداد المستعملة في عملية الفرز السريع، سنة 2001، في البيرو من أجل توضيح كيفية إجرائها:

يوم الانتخابات، تألفت الآلة الانتخابية في البيرو من 90780 مركزاً اقتراعياً. أولاً، يُقسم عدد المراكز الاقتراعية على عدد المراكز المطلوبة في العينة $(90780 \div 1020 = 89)$. يشير هذا العدد إلى اختيار مركز واحد من بين كل 89 مركزاً للتصويت. ثم يتم اختيار نقطة انطلاق عشوائية، عبر وضع 89 قصاصة من الورق مرقمة من 1 إلى 89 في وعاء، فيجري سحب إحدى الأوراق بطريقة عشوائية. حملت القصاصة التي تم اختيارها العدد 54. لذا، يدرج المركز الاقتراعي الذي يحمل الرقم 54 على اللائحة العشوائية، بصفته أول مركز لأخذ العينات؛ ثم يتم اختيار المراكز الاقتراعية التسعة والثمانين كلها، بعد اختيار المركز العيني الأول. وفقاً لذلك، حمل المركز الاقتراعي الثاني في العينة الرقم 143 على اللائحة (54 زائد 89). ويعاد هذا الإجراء حتى بلوغ حجم العينة الإجمالي: 1020.

لم يجب تنظيم لائحة المراكز الاقتراعية بطريقة عشوائية؟ تزيد هذه الاستراتيجية من صلاحية عملية الفرز السريع وموثوقيتها. ففي حال تنظيم اللائحة بحسب الحجم أو المنطقة أو المعايير الأخرى، قد تكون نتائج السحب الوحيد منحازة. لا يعتبر ذلك مشكلة كبيرة عادةً، غير أن التنظيم العشوائي هو تقنية توفر المزيد من الثقة بأن أرجحية اختيار كل مركز من العينة تساوي فرص اختيار أي مركز آخر.

عوامل التصحيح

لا بدّ أحياناً من تعديل عناصر مختلفة من منهجية الفرز السريع. تطبق هذه التعديلات على استقطاب المتطوعين وتدريبهم، وعلى عناصر أكثر تقنية ضمن الفرز السريع، بما في ذلك أخذ العينات. كما تستلزم حسابات العينات التي شرحناها أعلاه بعض التعديلات الإضافية، نظراً إلى الافتراض، أساساً، أنه سيجري تحديد كل مراكز أخذ العينات وسيتم تسليم المعطيات من كل مركز. ولكن لم تتمكن أية عملية فرز سريع واسعة النطاق أجرتها أية مجموعة مراقبة من نقل المعطيات من كل مركز معطيات فردي مذكور في العينة الأساسية.

في حالات الفرز السريع، من المهم التمييز بين عينة نظرية وعينة عملية. نفترض معظم المناقشات النظرية حول أخذ العينات أن معطيات المركز

التنظيم العشوائي هو تقنية توفر المزيد من الثقة بأن أرجحية اختيار كل مركز من العينة تساوي فرص اختيار أي مركز آخر.

لم تتمكن أية عملية فرز سريع واسعة النطاق أجرتها أية مجموعة مراقبة من نقل المعطيات من كل مركز معطيات فردي مذكور في العينة الأساسية.

المعنيّ ستصدر بفعالية مئة بالمئة، بعد أن يتمّ اختياره. غير أنّ هذا الافتراض لم يلقَ صدى واسعاً في أيّ فرز سريع، على صعيدٍ وطنيٍ رحب؛ ومردّد ذلك المزج بين العناصر، بأيّة طريقةٍ ممكنة، بما في ذلك الأخطاء التي يرتكبها المراقبون الخاضعون لتدريب غير ملائم، أو الأعطال التي تطرأ على أنظمة الاتصالات، أو تطورات اليوم الانتخابي غير المرتقبة. (يمنع، على سبيل المثال، المراقبون من دخول المراكز الاقتراعية، أو يحول الطقس العاصف دون تمكّن المراقبين من الوصول إلى الهاتف أو الإبلاغ عن المعطيات).

إبّان المعارك

الانتخابية الساخنة،
يمسي فقدان المعطيات
مسألة خطيرة للغاية.

تستطيع المنظّمات المدنيّة التي تجري عمليّة الفرز السريع، للمرّة الأولى، تسليم حوالي 75 بالمئة تقريباً من المعطيات الصادرة عن مراكز أخذ العينات، ضمن إطار زمني معقول يبلغ 3 ساعات. أما 25 بالمئة من العينة التي لم يتمّ التبليغ عنها (المعطيات المفقودة)، فقد تؤدي إلى مشاكل في تفسير المعطيات الأخرى. وبالتالي، تعتبر العينة العمليّة الصالحة للاستعمال أصغر دائماً من العينة النظرية. كما تكون هوامش الخطأ التي تنطبق على العينة العمليّة أكبر، بالضرورة، من تلك التي تمّ التخطيط لها.

إبّان المعارك الانتخابية الساخنة، يمسي فقدان المعطيات مسألة خطيرة للغاية. فضلاً عن أنّ هذه المعطيات المفقودة تتعدّى مجرد كونها عينة متنوّعة عشوائية من مجمل العينة. لطالما ارتفعت نسب فقدان المعطيات، عملياً، في المناطق النائية حيث تصعب إمكانية استرجاعها. وتجدر الإشارة إلى أنّ المعطيات المفقودة، إن لم تكن عشوائية أو تمثيلية، تمسي بالتالي مُحازرة. وفي حال افتقرت المعطيات المفقودة إلى الحياد، فإنّ الأمر نفسه ينطبق على العينة المتبقية.

ما هي الطريقة الفضلى لاستيعاب فكرة العجز عن استرجاع العينات كافة في اليوم الانتخابي؟ يجب إيجاد الحلّ في تصميم العينة الأساسي نفسه؛ كما يجب زيادة العينات وفقاً لهوامش نسبة الاسترداد المتوقعة.

على الإحصائي زيادة
العينات وفقاً لهوامش

نسبة الإستراداد
المتوقعة.

قد تتممّ مجموعة المراقبة الخبيرة بقدرةٍ على استرداد المعطيات بنسبة 80 بالمئة من مراكز العينات في العينة النظرية. في هذه الحالة، تعتبر العينة العمليّة أصغر بنسبة 20 بالمئة من العينة النظرية. أما الطريقة المباشرة لمواجهة هذه المشكلة المحتملة، فتتصّ على زيادة حجم العينة بنسبة 20 بالمئة، من خلال إضافة 20 بالمئة من مراكز العينات، بطريقةٍ عشوائية، على العينة التي تمّ احتسابها أولاً. قد تصلح هذه الاستراتيجية المباشرة في حال تمّ توزيع العجز في عملية استرداد العينة على السكان عشوائياً. غير أنّ التجارب السابقة تشير إلى أنّ العجز يكون عادةً موزّعاً، بنحو متفاوت، بين العاصمة والمناطق الحضرية الأخرى وبين المناطق الريفية. بطبيعة الحال، تظهر الصعوبة الأشدّ في المناطق النائية؛ من هنا، يجب أن يؤخذ هذا الأمر في الحسبان عند تصميم عنصر تصحيحيّ زائد من العينات. توضح الصورة 4-5 توزيع مخطط استرداد العينات النموذجي، وعنصر زيادة العينات التصحيحي. وكما تشير الصورة 4-5، قد يضع التصحيح الإضافي للاسترداد المتفاوت للعينات نصف العينات الزائدة على الأقل في المناطق الريفية.

الصورة 5-4:
نمط استرداد العينة
النموذجي وتوزيع زيادة
العينة الموصى به.

توزيع زيادة العينات	استرداد العينة	
15%	85%	العاصمة
25%	75%	المناطق الحضرية خارج العاصمة
35%	65%	المناطق النائية

أسئلة متكررة

أليس حجم العينة كبيراً للغاية في مطلق الأحوال، حيث أنها تؤخذ من المراكز الاقتراعية وليس من الناخبين المنفردين؟ إن صح ذلك، هل يستطيع المحللون معرفة حجم المعطيات بعد أخذ العينات؟

نعم، صحيح أن العينة كبيرة في الشكل الذي تؤخذ فيه، غير أن معرفة حجم المعطيات لا يحل محل المعطيات الحقيقية. فإن مجرد معرفة الحجم يمنح المعطيات "حجماً أكبر" في العينة الإجمالية. ما من طريقة لمعرفة إن كانت المعطيات الصادرة عن منطقة بعيدة عن العينة تمثل تلك التي تم استردادها من هذه العينة الثانوية بالذات. تشكل معرفة الحجم الوسيلة الأخيرة في الإحصائية الاستراتيجية التي لا يستحسن استعمالها إلا بعد نفاذ الخيارات الأخرى كافة.

تصحيح حجم المركز الاقتراعي

لا بدّ أحياناً من تصحيح هامش الخطأ في نتائج الفرز السريع بسبب اعتبارات عملية. فعلى سبيل المثال، يؤثر حجم المركز الاقتراعي - أي عدد الناخبين الإجمالي المتوقع في مركز الاقتراع - على هامش الخطأ. يعود هذا الأمر إلى الفرق بين عدد السكان المحدد وبين وحدة التحليل. تذكر أن حساب هامش الخطأ الأساسي قد اعتمد على عدد الناخبين المؤهلين للتصويت الإجمالي. وقد أجريت هذه العملية للتأكد من أن تصميم العينة قد التزم ببعض المبادئ الإحصائية. لكن بما أن مراكز الاقتراع هي وحدة التحليل المعتمدة، فمن المستحسن تعديل هامش الخطأ الذي يركز على عدد الناخبين في مراكز الاقتراع. في المثل السابق، تمّ تحديد ما يعادل 160 ناخباً لكل مركز اقتراعي. من المهمّ الأخذ بالاعتبار أن مراكز الاقتراع قد تكون مختلفة الأحجام. ففي حال حوت مراكز الاقتراع 200 ناخب، قد يؤدي ذلك إلى تقليص عدد المراكز المطلوبة في العينة. أما في حال كانت مراكز الاقتراع أكبر، أي تحوي 500 ناخب، فستكون المراكز اللازمة أقلّ بكثير.

يؤثر عدد المراكز
الاقتراعية وعدد
الناخبين في مركز
اقتراعي على هامش
الخطأ.

كما توضح الصورة 5-5، يؤثر عدد المراكز الاقتراعية وعدد الناخبين في مركز اقتراعي على هامش الخطأ. يعزى هذا

الأمر إلى دور حجم العينة في تحديد هامش الخطأ. تذكر أن معادلة هامش الخطأ هي:

(التغاير المفترض)★(القيمة Z في مستوى الثقة المختار)

$$\sqrt{n}$$

وفي الواقع، تؤثر الاختلافات في حجم المركز الاقتراعي أيضاً على العنصر 'n'.¹

الصورة 5-5:
حجم العينة وهامش
الخطأ

العينة	المراكز الاقتراعية		
	حجم المركز ناخبا 160	حجم المركز ناخب 200	حجم المركز ناخب 500
163,185	1,020	816	324
هامش الخطأ (مستوى ثقة بنسبة 95%)	±0.24	±3.43	±5.4
هامش الخطأ (مستوى ثقة بنسبة 99%)	±0.32	±4.5	±7.1

تجدر الإشارة إلى أن هامش الخطأ منوطٌ بعدد مراكز الاقتراع في العينة. في حال كانت المراكز الاقتراعية كبيرة، يلزم عددٌ أقلّ منها لإنشاء العينة المطلوبة التي يبلغ حجمها 163185 ناخباً. ويفوق هامش الخطأ الذي يحتسب للمراكز الاقتراعية ذلك الذي يحتسب لعينة الناخبين. فيقع هامش الخطأ الخاص بعمليات الفرز السريع بين هامش الخطأ الأدنى والأقصى.

كلّما انخفض عدد

المراكز الاقتراعية

المطلوب لتكوين عينة

من الناخبين كلّما زاد

هامش الخطأ.

بفضل اقتفاء أثر التبدلات في هامش الخطأ بالنسبة لأحجام مختلفة من مراكز الاقتراع، ينخفض عدد المراكز الاقتراعية المطلوب لتكوين عينة من الناخبين كلّما زاد هامش الخطأ. توضح الصورة 5-6 العلاقة بين حجم المركز الاقتراعي وبين هامش الخطأ.

الصورة 5-6:
حجم المركز الاقتراعي
وهامش الخطأ

الناخبون/ المركز	عدد المراكز اللازمة لأخذ العينة	هامش الخطأ (مستويات الثقة)	
		%95	%99
150	1088	±2.97	±3.91
200	816	±3.43	±4.52
250	653	±3.84	±5.05
300	544	±4.20	±5.53
350	466	±4.54	±5.97
400	408	±4.85	±6.39
450	363	±5.14	±6.77
500	327	±5.42	±7.13

يزيد هامش الخطأ مع زيادة حجم المركز الاقتراعي. غير أن التأثير العام لحجم المركز الاقتراعي على هامش الخطأ يتدنى لدى ازديادهما معاً.



الصورة 5-7:
حجم المركز الاقتراعي
وهامش الخطأ

التصحيح وفقاً لمعدل المشاركة في الانتخابات

عندما تكون المعركة الانتخابية حامية، يجب أن يهتم المحللون أيضاً بنسبة مشاركة الناخبين في الانتخابات. وبالرغم من نجاح المراقبين في استرجاع المعطيات من كل مركز بين 1020 مركزاً اقتراعياً مدرجاً في العينة النظرية، فإن المشاركة الدنيا للناخبين في العملية الانتخابية تفترض أن العينة ستتضمن عدداً أقل من الأصوات التي كانت ستشملها لو كانت نسبة المشاركة مرتفعة. وقد ارتكزت الحسابات في الأصل على توقع مشاركة 160 ناخباً في كل مركز اقتراعي. ولكن في حال بلغت المشاركة نسبة 70 بالمئة، قد يكون عدد الناخبين في كل مركز اقتراعي 112 ناخباً ليس إلا. وفي حال تكرار هذا النمط في المراكز الاقتراعية جميعها، قد يتضمن الفرز حينذاك 114240 ناخباً وحسب، أي مشاركة خجولة تبلغ 50000 من أصل 163185 ناخباً مطلوباً لبلوغ هامش خطأ بنسبة 0.3 بالمئة ومستوى ثقة بنسبة 99 بالمئة.

عندما تكون المعركة الانتخابية حامية، يجب أن يهتم المحللون أيضاً بنسبة مشاركة الناخبين في الانتخابات.

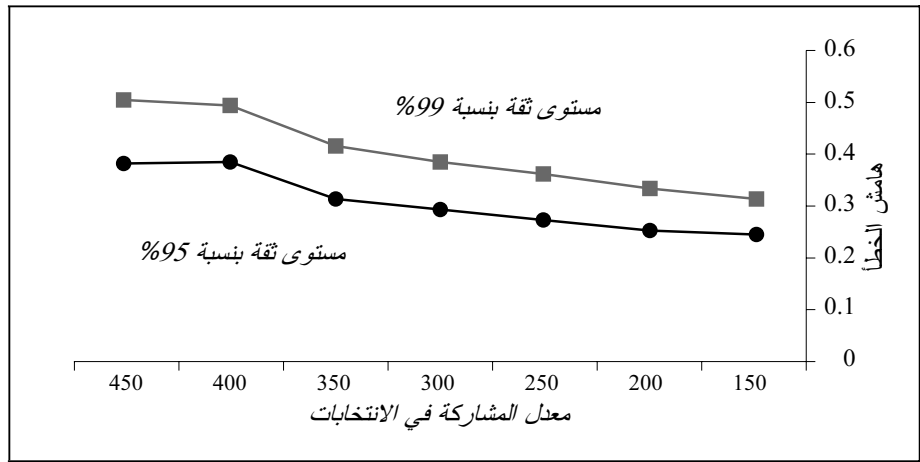
هامش الخطأ (مستويات الثقة)		عدد الناخبين، الأصوات	نسبة مشاركة الناخبين
%99	%95		
0.31±	0.24±	163185	العينة المطلوبة (مشاركة بنسبة 100%)
0.33±	0.25±	146867	%90
0.36±	0.27±	130548	%80
0.38±	0.29±	114230	%70
0.41±	0.31±	97911	%60
0.49±	0.38±	81593	%50

الصورة 5-8:
نسبة مشاركة الناخبين
وهامش الخطأ

بالتالي، تدعو استراتيجية تفسير حذرة للمعطيات إلى إعادة احتساب هامش الخطأ، بناءً على عدد الأصوات الفعلية التي جرى فرزها. توضح الصورة 5-8 هذا المثل.

كما يظهر الجدول، كلما تقلصت نسبة مشاركة الناخبين، كلما زاد هامش الخطأ. فإذا تجاوزت نسبة المشاركة 60 بالمئة، سيرتفع هامش الخطأ بنسبة 0.02 بالمئة تقريباً لكل هبوط بنسبة 100 بالمئة في معدل المشاركة في الانتخابات. لدى اقتراب هذا المعدل من نسبة 50 بالمئة، تصبح الزيادة في هامش الخطأ أعلى بكثير. توضح الصورة 5-9 رسماً بيانياً حول ارتفاع هامش الخطأ، بالتوازي مع انخفاض معدل المشاركة في الانتخابات.

الصورة 5-9:
هامش الخطأ ومعدل
المشاركة في الانتخابات



حدّد هذا الفصل المبادئ الإحصائية العامة التي ترسي قواعد عمليّات الفرز السريع من أجل جمهور واسع، كما أنه أوجز الأساسات الإحصائية لمنهجية الفرز السريع. على المراقبين فهم هذه المنهجية، لا سيّما مفاهيم الموثوقية والصلاحية، فضلاً عن فهم سبب التزام العينة بمعيار العشوائية. تتشكل هذه المعرفة عنصراً حيوياً في تصميم استمارات مراقبة فعّالة، ويمكن الوثوق بها، بالإضافة إلى البرامج التدريبيّة. فضلاً عن أنّها تبرز أهمية الإعداد لاسترجاع المعطيات من أنحاء البلاد كافة، وحتى من أبعد المناطق.

أخيراً، تطرّق هذا الفصل إلى مسائل أكثر تقنيّة حول كيفية احتساب أحجام العينة وكيفية تأثير بعض الأمور كمستويات الثقة، وهوامش الخطأ، وتغاير السكان، وتجانسهم، على العينة. ينشد معظم المراقبين خدمات إحصائيّ خبير من أجل إعداد عينة وأخذها، وتحليل المعطيات في اليوم الانتخابي. على المجموعات المدنيّة أن تدرك أنّ الفرز السريع يتعلّق بتطبيق مبادئ إحصائية على ظروف عمليّة وفريدة من نوعها، عندما يستحيل إرضاء افتراضات الكتب التعليميّة. ولهذا السبب، يوضح هذا الفصل ما هي العناصر التصحيحية الأكثر شيوعاً التي يجب أخذها في الحسبان، عندما يدرس المحللون تفسير المعطيات التي يحصلون عليها بنجاح يوم الانتخابات.

تذكير



يفهم المسؤولون غير الإحصائيين بسهولة المبادئ العامة التي تسيطر على عمليات الفرز السريع؛ وتتعدد الأسباب الوجيهة التي تناشد الموظفين الرئيسيين في مجموعات المراقبة التآلف وهذه المبادئ:

1. إن إدراك أهمية ضمان صلابة معطيات الفرز السريع يسهل اتخاذ القرارات بشأن تصميم الفرز السريع، ويساعد الموظفين على تطوير استثمارات مراقبة فعالة وبرامج تدريبية.

2. إن الموظفين الذين يقدرّون العلاقة بين العينة، والسكان، وأهمية متطلبات العشوائية في تأمين سلامة هذه العلاقة، مدعون إلى إنشاء شبكة صلبة من المنطوقين تغطي حتى أبعد المراكز الاقتراعية.

يترتب على المجموعات الاستعانة بإحصائيين خبراء في مجال عمليات الفرز السريع، من أجل تنفيذ مهمات معقدة تقنياً، كأعداد عينة وتحليل نتائج عمليات الفرز السريع. أما التمتع بالخبرة في مجال عمليات الفرز السريع في أنحاء عدة من العالم فيسلط الضوء على نقاط عدة:

1. إن وحدة تحليل الفرز السريع هي المركز الاقتراعي. لا يمكن المباشرة بعمليات أخذ العينات قبل توافر لائحة دقيقة وشاملة بمراكز الاقتراع- أي إطار أخذ العينات.

2. تستعمل عمليات الفرز السريع بصورة دائمة عينات الأرجحية (مثلاً، العينات العشوائية العامة أو العينات العشوائية الطبقيّة) من أجل الحصول على نتائج تمثل السكان كافة.

3. لا تستطيع مجموعات المراقبة التي تجري عمليات الفرز السريع استرجاع 100 بالمئة من معطيات العينة. فعلى المحللين الاستعداد لهذا الأمر المحتم؛ أما الحلّ الذي يمكن اعتماده في تصميم العينة الأصلي، فيتمثل بزيادة العينة وفق هوامش نسبة الاسترداد المتوقعة.

4. على المحللين أن يأخذوا بعين الاعتبار أيضاً العناصر التصحيحية لدى إعداد العينة. أما العناصر الأهمّ فهي تلك التي تأخذ في الحسبان الاختلافات في (أ) معدل مشاركة الناخب في الانتخابات، و(ب) عدد الناخبين في وحدة التحليل الأساسية، أي المركز الاقتراعي.