UN SONDAGE SENSIBLE

Question
Avez-vous déjà pris de la drogue ?

Lancer de dé

Si résultat pair → DIRE LA VÉRITÉ

Si résultat impair → RELANCER LE DÉ

Si résultat pair → RÉPONDRE OUI

Si résultat impair → RÉPONDRE NON

Comment répondre à un sondage sensible en dissimulant sa réponse de façon efficace et résistante aux attaques ? Grâce au mécanisme de réponse randomisé proposé par Stanley Warner en 1965. Demandons à 1000 individus s'ils ont déjà pris de la drogue. Selon le résultat d'un jet de dé (pair ou impair), ils disent la vérité ou bien répondent au hasard en fonction d'un second jet de dé. La probabilité de répondre honnêtement *versus* aléatoirement (ici, 1/2) est fonction du paramètre de vie privée ε. Résultat : 500 interrogés répondent au hasard, et parmi eux, 250 prétendent avoir pris de la drogue, sans lien avec la vérité. Si, sur les 1000 interrogés, 450 individus répondent « oui », seuls 200 font partie de la moitié (500 individus) qui dit la vérité. Sur l'ensemble, en supposant que le groupe des interrogés ayant répondu au hasard suit la même distribution que le groupe ayant répondu honnêtement, c'est donc à peu près le double (400) qui a déjà pris de la drogue.

Ce développement récent d'outils a accompagné la mise en place de la confidentialité différentielle au sein d'applications et de services de la vie quotidienne. Cependant, son utilisation par les acteurs du numérique est pour l'instant concentrée sur la collecte de statistiques ou bien la publication d'informations de nature statistique.

ET DANS LA VRAIE VIE ?

En 2016, lors de sa conférence annuelle WWDC, Apple introduisait l'utilisation de la confidentialité différentielle pour, entre autres, la collecte de statistiques sur l'utilisation des émojis dans les messages privés. Depuis, l'idée, auparavant cantonnée au milieu académique, s'est diffusée dans les médias qui, même spécialisés, touchent un plus grand public. Avant cela, en 2014, Google avait mis en place un outil nommé Rappor afin d'analyser l'utilisation de son navigateur Chrome. Grâce à cet algorithme, fondé sur un mécanisme proposé en 1965 par le statisticien Stanley Warner, un individu peut répondre à un sondage sensible en conservant la possibilité de nier de façon plausible sa réponse, tout en offrant des garanties sur la quantité de bruit une fois l'ensemble des réponses agrégées. Le mécanisme en jeu est le suivant : étant donné une question sensible dont la réponse est binaire (« oui » ou « non »), l'individu choisit aléatoirement, par exemple à l'aide d'un jet de dé, de répondre la vérité (avec une certaine probabilité qui dépend du bruit, c'est-à-dire au paramètre ε), ou de répondre « oui » ou « non » aléatoirement à l'aide d'un second jet de dé *(voir l'encadré ci-dessus)*. Les résultats individuels sont agrégés dans une structure probabiliste qui donne ensuite accès à des statistiques globales correctes tout en respectant la confidentialité différentielle.

Plus récemment, des jeux de données d'importance ont été anonymisés par des algorithmes satisfaisant la confidentialité différentielle. En 2020, le bureau du recensement de la population aux États-Unis (Census) l'a adoptée comme modèle de vie privée pour la publication de ses statistiques. En plus des nombreuses valeurs agrégées publiées, le Census distribue des données synthétiques générées aléatoirement à partir des distributions anonymisées. La confidentialité différentielle a aussi été utilisée pour publier le