

Section 1. Enquêtes probabilistes et non probabilistes

Cette section explique la différence entre les enquêtes probabilistes et non probabilistes ainsi que l'approche adoptée par la méthodologie MOWIP pour ses enquêtes.

Enquêtes probabilistes

Les enquêtes probabilistes ou enquêtes par sondage désignent des enquêtes qui s'appuient sur un échantillon représentatif d'une population globale afin d'étudier différents sujets : attributs, comportements, opinions, etc. Les enquêtes probabilistes sont utilisées lorsqu'il n'est pas possible d'interroger chaque individu d'une population comme lors d'un recensement. Les sondages d'opinion politique sont un exemple d'enquête probabiliste où un échantillon représentatif de la population globale est tiré au sort. Cette forme d'enquête est très répandue en raison des ressources considérables (financières, humaines, mais aussi en termes de temps) qu'exigerait l'interview de chaque individu de la population visée. Pour reprendre l'exemple de l'opinion politique, une organisation ou institution indépendante disposant de ressources limitées pourrait difficilement interroger toute la population des États-Unis (plus de 300 millions de personnes). Un échantillon représentatif est donc sélectionné selon des méthodes scientifiques qui permettent de fournir des informations utiles sur l'ensemble de la population. Les résultats de l'enquête peuvent ensuite être généralisés à l'ensemble de la population étudiée. Pour comprendre l'échantillonnage probabiliste, il faut savoir que tous les individus de la population ont les mêmes chances d'être sélectionnés pour l'étude et que l'échantillon sera choisi au hasard. Par exemple, si l'on souhaite étudier l'ensemble d'une force de police dans un pays donné en s'appuyant sur un échantillonnage probabiliste, il faut sélectionner aléatoirement un échantillon représentatif d'agent·e·s de police qui donne à chaque agent·e de police du pays concerné la même chance d'être sélectionné·e pour l'enquête.

Enquêtes non probabilistes

Les enquêtes non probabilistes désignent des enquêtes dans lesquelles les individus d'une population n'ont pas tous les mêmes chances d'être sélectionnés. Si certains membres de la population sont sciemment écartés de la stratégie d'échantillonnage, alors la méthode utilisée est celle d'une enquête non probabiliste. Par exemple, si une équipe de recherche décide d'envoyer des courriels pour inviter à participer à une enquête d'opinion politique, elle écarte sciemment les personnes qui n'ont pas d'adresse électronique. Par conséquent, les résultats ne pourront pas être généralisés à l'ensemble de la population. Les enquêtes non probabilistes sont généralement utilisées pour étudier une population spécifique en interrogeant un pourcentage déterminé des individus d'un groupe donné.

L'approche de la méthodologie MOWIP

En règle générale, la méthodologie MOWIP recourt à des méthodes d'échantillonnage non probabiliste en connaissance de cause et pour des raisons très précises.

Premièrement, il s'agit de s'assurer que l'on interroge un nombre de femmes suffisamment important dans le cadre de l'échantillon. La méthodologie MOWIP a pour but d'évaluer l'existence éventuelle d'obstacles à la participation significative des femmes aux opérations de paix. À ce titre, il faut s'assurer que l'on interroge suffisamment de femmes dans le cadre des enquêtes auprès des membres du personnel déployé·e·s et non déployé·e·s. Pour ce faire, on utilise une stratégie appelée « échantillonnage par quotas », qui consiste à diviser l'ensemble de la population des forces de sécurité en sous-groupes (femmes et hommes, personnel déployé et non déployé, par exemple), puis à interroger un certain nombre de femmes et de membres du personnel déployé. Si le choix de l'échantillon reposait sur une méthode probabiliste, il y aurait une grande probabilité que l'échantillon comprenne moins de femmes et de personnel déployé dans les institutions comptant moins de femmes ou moins de personnel déployé. Dès lors, pour pouvoir interroger un nombre suffisant de personnes entrant dans ces catégories, la méthodologie établit un quota de participantes et de membres du personnel déployé et elle recommande de suréchantillonner les personnes concernées. Cela dit, si l'on a la certitude de pouvoir échantillonner un nombre suffisant de femmes et de membres du personnel déployé dans un pays donné, c'est à la méthode probabiliste qu'il convient de recourir. Les stratégies d'échantillonnage sont détaillées ci-après.

Section 2 : Marge d'erreur, niveau de confiance et intervalle de confiance

Cette section explique comment la marge d'erreur, les niveaux de confiance et les intervalles de confiance sont calculés dans le cadre de la méthodologie MOWIP.

Les résultats des enquêtes relevant de la méthodologie MOWIP visent une **marge d'erreur** inférieure ou égale à 5 % et un **niveau de confiance** supérieur ou égal à 95 %.

La **marge d'erreur** désigne la différence maximale possible, exprimée en pourcentage, entre les résultats de l'étude (donnés par l'échantillon) et la **population** réelle. La **population** désigne l'ensemble du groupe d'individus auquel s'appliquent les résultats de l'étude. Aux fins de la méthodologie MOWIP, la population désigne tou·te·s les membres de l'organisation de sécurité évaluée. Comme il n'est pas possible (ou nécessaire) d'interroger toute la population, on interroge seulement une partie de celle-ci. L'**échantillon** désigne ainsi les membres de l'institution de sécurité qui participent à l'étude en qualité de répondant·e·s.

L'**intervalle de confiance** désigne une fourchette de valeurs ou un intervalle numérique à l'intérieur duquel se situe le véritable paramètre de la population. Autrement dit, en fonction de l'intervalle de confiance, l'équipe de recherche est en mesure d'estimer quelle pourrait être la valeur réelle s'appliquant à la population (une estimation ponctuelle située dans la fourchette de l'intervalle de confiance). Le **niveau de confiance** désigne la probabilité que l'intervalle de confiance inclue le paramètre véritable de la population. Par exemple, un niveau de confiance de 95 % signifie que si l'étude est menée à plusieurs reprises selon les mêmes techniques, dans 95 % des cas, la véritable réponse se situera à l'intérieur de l'intervalle de confiance. Celui-ci reflète le degré d'incertitude des statistiques présentées dans l'étude.

La marge d'erreur fournit la plage de valeurs à ajouter ou à déduire d'une estimation ponctuelle, ce qui nous donne la fourchette des intervalles de confiance. Par exemple, un niveau de confiance de 95 % avec une estimation ponctuelle de 45 et une marge d'erreur de 5 % signifie que les résultats seront compris, 95 % du temps, dans une fourchette de 5 points de pourcentage autour de 45. Autrement dit, la marge d'erreur représente la différence maximale entre la valeur réelle et l'estimation ; et l'intervalle de confiance est l'estimation plus et moins la marge d'erreur.

Si l'on reprend l'exemple utilisé pour la marge d'erreur, cela signifie qu'en réalisant une fois l'étude, l'équipe de recherche est sûre à 95 % que le résultat obtenu se situe dans une fourchette de 5 % par rapport à la réponse qu'elle aurait obtenue en interrogeant l'ensemble du personnel. Cependant, si elle réalisait cette étude 20 fois avec des échantillons différents, il faudrait s'attendre à obtenir une fois un résultat « faux », c'est-à-dire que la réponse obtenue à partir de l'échantillon s'écarterait de plus de 5 % de la réponse qu'elle aurait obtenue avec l'ensemble de la population.

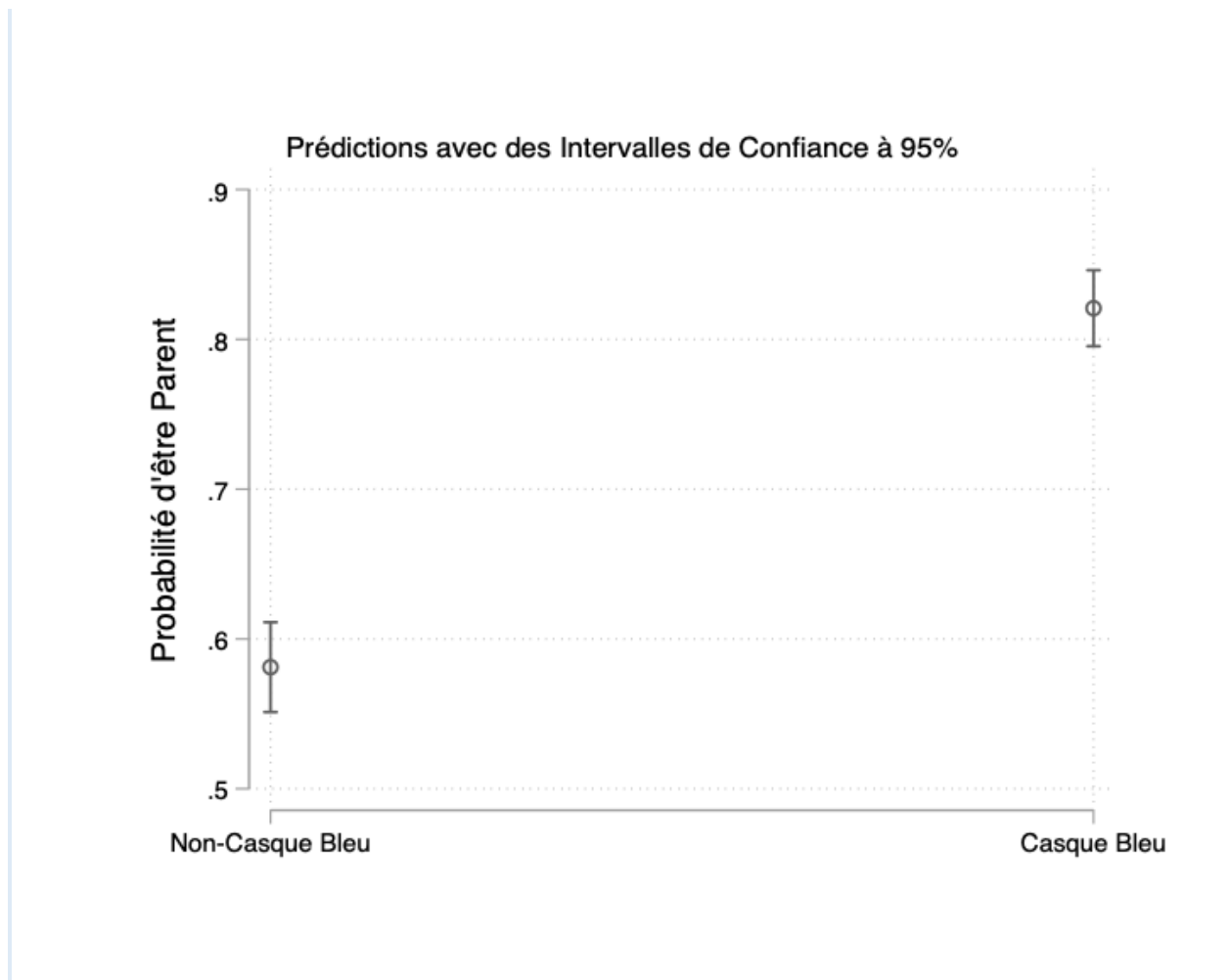
Exemple

L'étude du personnel des forces de sécurité au Ghana, en Uruguay, en Zambie et au Sénégal montre que la probabilité que les Casques bleus aient des enfants est de 82 %, avec un niveau de confiance de 95 % et une marge d'erreur de 2,5 %.

Si l'étude était répétée indéfiniment selon les mêmes techniques, chaque résultat devrait se situer à moins de 2,5 % du résultat indiqué (82 %) dans 95 % des cas. Autrement dit, dans 95 % des cas, le résultat devrait se situer entre :

- $82 - 2,5 = 79,5 \%$
- et
- $82 + 2,5 = 84,5 \%$

Le graphique ci-dessous illustre la probabilité que les Casques bleus aient des enfants par rapport aux autres membres des forces de sécurité. Il présente l'estimation ponctuelle (le cercle) et les intervalles de confiance (les lignes au-dessus et au-dessous de l'estimation ponctuelle). La probabilité que les Casques bleus aient des enfants se situe entre 79,5 % et 84,5 % et la meilleure estimation est de 82 % (le cercle). Avec un niveau de confiance de 95 %, on peut dire que si l'on sélectionne des échantillons du personnel des forces de sécurité selon les mêmes techniques, la probabilité que les Casques bleus aient des enfants se situerait dans 95 % des cas entre 79,5 % et 84,5 %. De même, la probabilité de ne pas être Casque bleu et de ne pas avoir d'enfant se situerait entre 55 % et 61 %, la meilleure estimation étant de 58 %.



Calcul de l'intervalle de confiance et de la marge d'erreur

Pour calculer l'**intervalle de confiance**, l'équipe de recherche doit d'abord calculer les « degrés de liberté » et déterminer le « niveau de confiance ». Puis, pour calculer la marge d'erreur, elle doit calculer l'« erreur standard » et la « valeur critique ».

Les **degrés de liberté** désignent le nombre d'informations indépendantes incluses dans une estimation. Pour calculer ce nombre, l'équipe doit soustraire 1 à la taille de l'échantillon, c'est-à-dire au nombre d'éléments/répondants de l'étude. Par exemple, si la taille de l'échantillon est de 380, il y aura : $380 - 1 = 379$ degrés de liberté.

Le **niveau de confiance** désigne la probabilité (que l'on peut exprimer en pourcentage) que la valeur réelle de la population se situe dans une fourchette donnée.

Après avoir calculé les degrés de liberté et spécifié le niveau de confiance, il faut soustraire le **niveau de confiance** de 1, puis diviser le chiffre obtenu par deux pour déterminer l'alpha (α). Par exemple, si le niveau de confiance est de 95 %, alors $\alpha = (1 - 0,95)/2 = 0,025$.

L'équipe de recherche va alors consulter une table appelée « table de distribution T » pour trouver une valeur appelée la « valeur critique » ou « valeur T » pour 379 degrés de liberté (dl) et un α de 0,025, comme on le voit ci-dessous. La valeur T représente la différence entre un point de données et la moyenne. Elle permet de savoir si une valeur (l'âge du premier déploiement des Casques bleus, par exemple) est très éloignée de la moyenne (l'âge moyen du déploiement des Casques bleus au sein de l'institution). Les valeurs T sont présentées dans une table de distribution T qui est universelle. Comme le montre la figure 1, pour 379 degrés de liberté et un α de 0,025, la valeur critique s'élève à **1,966**.

Figure 1 : Comment utiliser la table de distribution T

	Area in right tail = 0.25	Area in right tail = 0.20	Area in right tail = 0.15	Area in right tail = 0.10	Area in right tail = 0.05	Area in right tail = 0.025	Area in right tail = 0.02	Area in right tail = 0.01	Area in right tail = 0.005	Area in right tail = 0.0025	Area in right tail = 0.001	Area in right tail = 0.0005
DF	t-score	t-score	t-score	t-score	t-score	t-score	t-score	t-score	t-score	t-score	t-score	t-score
359	0.675	0.843	1.038	1.284	1.649	1.967	2.061	2.337	2.590	2.824	3.113	3.318
360	0.675	0.843	1.038	1.284	1.649	1.967	2.061	2.337	2.590	2.824	3.113	3.318
361	0.675	0.843	1.038	1.284	1.649	1.967	2.061	2.337	2.590	2.824	3.113	3.318
362	0.675	0.843	1.038	1.284	1.649	1.967	2.061	2.337	2.589	2.824	3.113	3.318
363	0.675	0.843	1.038	1.284	1.649	1.967	2.061	2.337	2.589	2.824	3.113	3.318
364	0.675	0.843	1.038	1.284	1.649	1.967	2.061	2.337	2.589	2.824	3.113	3.317
365	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.337	2.589	2.824	3.113	3.317
366	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.337	2.589	2.824	3.113	3.317
367	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.337	2.589	2.824	3.113	3.317
368	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.337	2.589	2.824	3.113	3.317
369	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
370	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
371	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
372	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
373	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
374	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
375	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
376	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
377	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.317
378	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.316
379	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.316
380	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.824	3.112	3.316
381	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.112	3.316
382	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.112	3.316
383	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.112	3.316
384	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.112	3.316
385	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.112	3.316
386	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.111	3.316
387	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.111	3.316
388	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.111	3.316
389	0.675	0.843	1.038	1.284	1.649	1.966	2.061	2.336	2.589	2.823	3.111	3.316

L'erreur standard, ou erreur type, est une mesure de dispersion qui renseigne l'équipe de recherche sur la distribution des données. Plus elle est élevée et plus les données sont hétérogènes ou dispersées. L'erreur standard s'obtient en divisant l'« écart-type » par la racine carrée de la taille de l'échantillon. L'écart-type s'obtient quant à lui en calculant la moyenne de l'échantillon, en soustrayant chaque valeur de la moyenne, puis en divisant ce chiffre par les degrés de liberté.

Calcul de l'écart-type de l'échantillon (S)

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

- \sum = somme
- X = chaque valeur de l'échantillon
- \bar{x} = moyenne de la population
- n = nombre de valeurs dans l'échantillon

Une fois que l'on a calculé l'erreur standard et la valeur critique, on peut les multiplier ensemble pour obtenir la **marge d'erreur**.

Marge d'erreur = erreur standard x valeur critique

Une fois la **marge d'erreur** connue, on peut calculer l'**intervalle de confiance**, qui correspond simplement à la fourchette de valeurs que l'on obtient en prenant la moyenne de l'échantillon plus ou moins la marge d'erreur.

Intervalle de confiance = moyenne de l'échantillon \pm marge d'erreur

Exemple

On veut connaître la durée moyenne, en mois, du déploiement des Casques bleus dont il était question dans notre étude sur le personnel des forces de sécurité au Ghana, en Uruguay, en Zambie et au Sénégal. En moyenne, les 876 personnes ayant répondu à l'enquête ont été déployées pendant 19 mois. C'est donc la moyenne de l'échantillon. Nous allons nous en servir pour estimer quelle serait la moyenne de la population (c'est-à-dire la durée moyenne du déploiement que l'on obtiendrait en échantillonnant l'ensemble du personnel). L'écart-type est de 13 et le niveau de confiance choisi s'élève à 95 %, ce qui correspond à une valeur critique de 1,960 dans la table de distribution T.

On calcule d'abord l'erreur standard en divisant l'écart-type (13) par la racine carrée de la taille de l'échantillon (876).

$$13 / \sqrt{(876)} = 0,439$$

Ensuite, on calcule la marge d'erreur en multipliant l'erreur standard par la valeur critique, qui est de 1,966 (figure 1).

$$0,439 \times 1,960 = \mathbf{0,860} \cong \mathbf{1}$$

Comme on l'a vu, la marge d'erreur nous donne la fourchette de valeurs **au-dessus et en dessous** de la moyenne de l'échantillon. Cette fourchette correspond à l'intervalle de confiance. La limite supérieure est égale à la moyenne de l'échantillon plus la marge d'erreur, et la limite inférieure à la moyenne de l'échantillon moins la marge d'erreur :

Limite supérieure : $19 + 1 = 20$

Limite inférieure : $19 - 1 = 18$

Cela nous donne un intervalle de confiance compris entre 18 et 20. Ce qui signifie que si l'on recommence l'enquête plusieurs fois, on peut s'attendre à ce que la moyenne de l'échantillon

(la durée moyenne du déploiement des personnes interrogées) se situe quelque part entre 18 et 20 mois dans 95 % des cas.

Section 3 : Taille de l'échantillon

Cette section explique comment calculer la taille de l'échantillon dans le cadre de la méthodologie MOWIP.

Pourquoi la taille de l'échantillon est-elle importante ?

La méthodologie MOWIP recommande de s'appuyer sur un échantillon d'au moins 380 répondant·e·s pour pouvoir tirer des conclusions concernant les groupes visés. Dans le cadre d'une analyse scientifique des institutions de sécurité, pour pouvoir recommander d'adopter ou réviser des politiques et généraliser les conclusions d'une étude à l'ensemble de l'institution de sécurité concernée, il faut enquêter auprès d'un nombre suffisamment élevé de membres du personnel appartenant à différents groupes (hommes et femmes, différents grades, personnel déployé ou non, divisions et unités techniques variables, etc.). Comme on l'a vu dans la section précédente, la taille de l'échantillon joue un rôle important dans le calcul des différentes composantes de l'analyse, telles que la marge d'erreur, l'écart-type ou l'intervalle de confiance. Plus l'échantillon est important, plus l'incertitude des mesures diminue et plus les résultats seront fiables, c'est-à-dire que les résultats se rapprocheront de ce qu'on aurait trouvé dans la population réelle. Les figures 2 et 3 ci-dessous montrent que lorsque la taille de l'échantillon augmente, l'erreur standard et la marge d'erreur diminuent toutes les deux. Par conséquent, pour garantir une marge d'erreur de 5 % (valeur limite pour la méthodologie MOWIP), il faut disposer d'un échantillon composé d'un nombre suffisant de répondant·e·s.

Figure 2. Distribution de l'erreur standard par rapport à la taille de l'échantillon

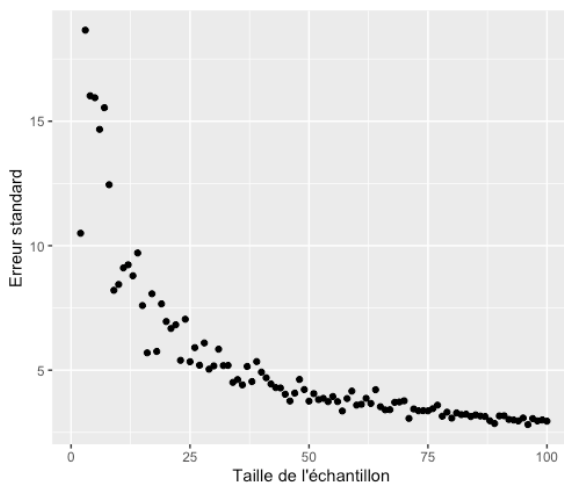
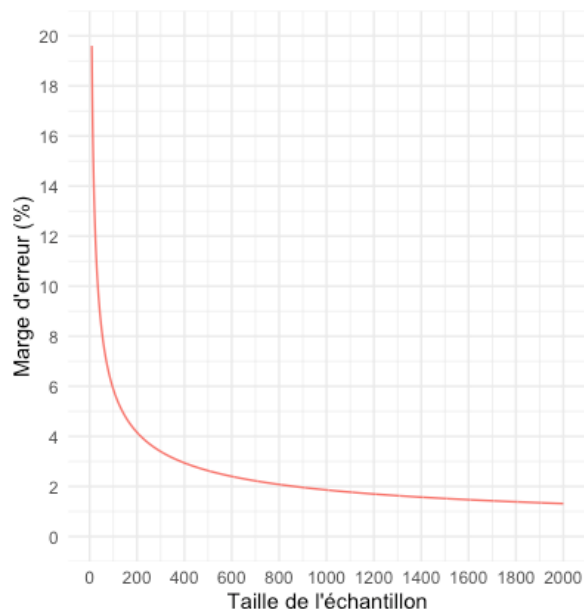


Figure 3. Distribution de la marge d'erreur par rapport à la taille de l'échantillon (intervalle de confiance de 95 %)



Variation des marges d'erreur en fonction de différentes tailles d'échantillon (hommes et femmes, femmes déployées, hommes déployés, femmes non déployées et hommes non déployés)

Les figures 4 à 6 illustrent la marge d'erreur selon différentes tailles d'échantillon. Comme on l'a vu, plus l'échantillon est fourni, plus la marge d'erreur diminue. La figure 4 montre qu'avec un échantillon de 380 personnes – la taille recommandée par la méthodologie MOWIP – la marge d'erreur est à peu près égale à 3 %. La figure 5 montre qu'avec un échantillon de 190 – la taille recommandée par la méthodologie MOWIP pour un échantillon hommes/femmes et un échantillon personnel déployé/non déployé – la marge d'erreur est d'environ 4 %. Enfin, la figure 6 montre qu'avec un échantillon de 95 – la taille recommandée par la méthodologie MOWIP pour un échantillon femmes déployées/hommes déployés – la marge d'erreur est d'environ 6 %.

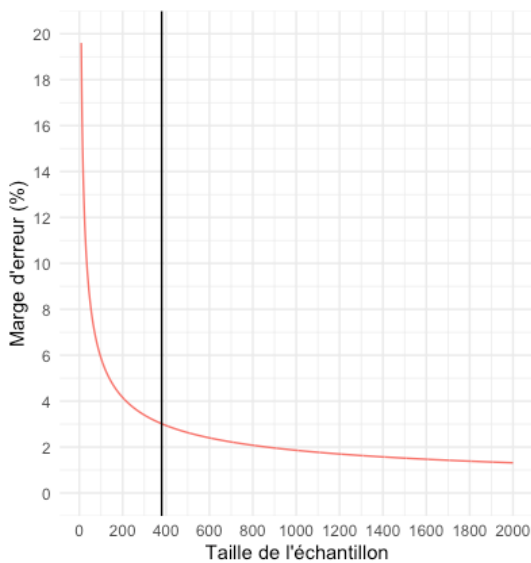


Figure 4. Marge d'erreur avec un échantillon de 380

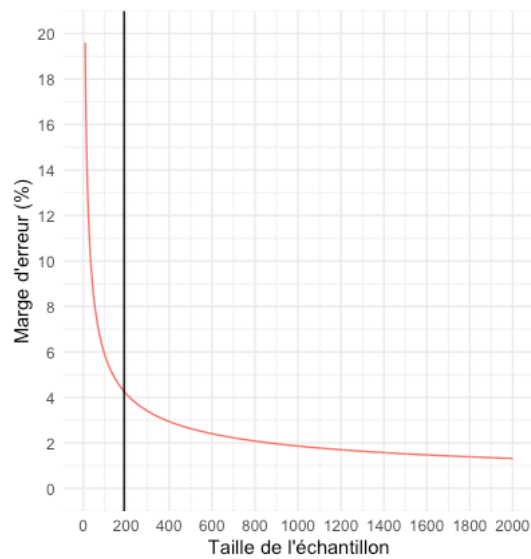


Figure 5. Marge d'erreur avec un échantillon de 190 (hommes/femmes)

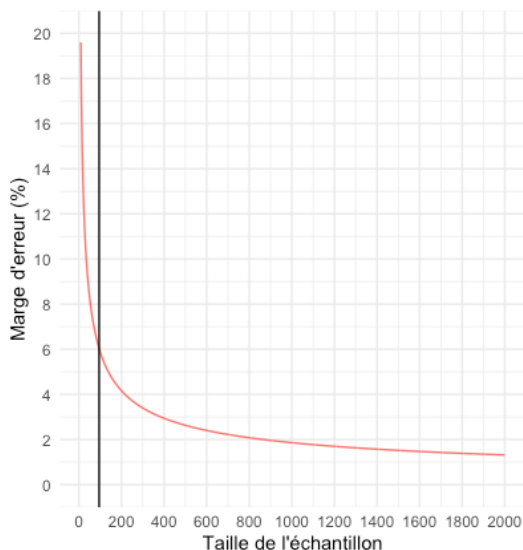


Figure 6. Marge d'erreur avec un échantillon de 95
(hommes/femmes déployé-e-s)

Variation des tailles d'échantillon en fonction de l'intervalle de confiance, de la marge d'erreur et de la taille de la population

La taille de l'échantillon nécessaire peut être déterminée à partir de l'intervalle de confiance et de la marge d'erreur que l'on souhaite. Ces deux variables établissent la part d'incertitude des mesures. Si l'équipe de recherche souhaite réduire le degré d'incertitude de ses résultats, elle doit augmenter ses niveaux de confiance et diminuer ses marges d'erreur, ce qui demande un échantillon plus important.

Cela dit, un autre facteur permet de déterminer la taille de l'échantillon nécessaire en fonction du niveau de confiance et de la marge d'erreur que l'on souhaite : **la taille de la population**. Plus la population est importante et plus il faudra choisir un échantillon important pour parvenir aux mêmes niveaux de certitude. Les figures 7 à 15 illustrent la distribution des tailles d'échantillon nécessaires pour différentes valeurs d'intervalle de confiance et de marge d'erreur en fonction de la taille de la population.

Les figures 7 à 9 illustrent la variation des tailles d'échantillon nécessaires selon la taille de la population pour un intervalle de confiance de 99 % et des marges d'erreur de 10 %, 5 % et 3 % respectivement. L'échantillon nécessaire sera plus important si l'on opte pour un niveau de confiance de 99 % et une marge d'erreur de 3 %. Pour une institution de sécurité comptant 5000 employé·e·s, par exemple, si l'on opte pour un niveau de confiance de 99 % et une marge d'erreur de 5 % afin d'évaluer l'institution, il faudra interroger environ 600 personnes, comme le montre la figure 8 (en traçant une ligne verticale partant du chiffre 5000 sur l'axe de la taille de population pour rejoindre la courbe en rouge, puis en rejoignant l'axe de la taille d'échantillon en suivant une ligne horizontale).

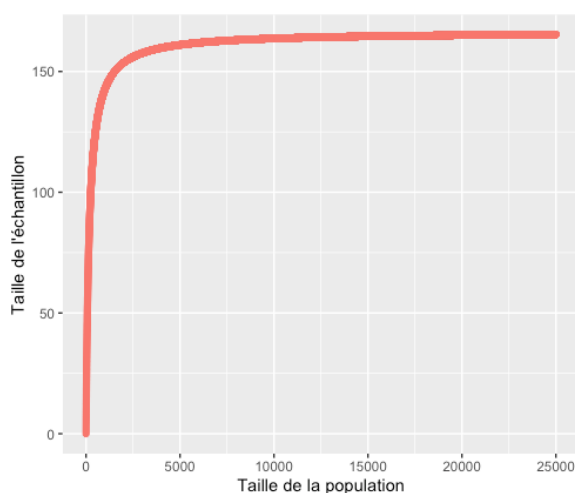


Figure 7. Taille d'échantillon pour un intervalle de confiance de 99 % et une marge d'erreur de 10 %

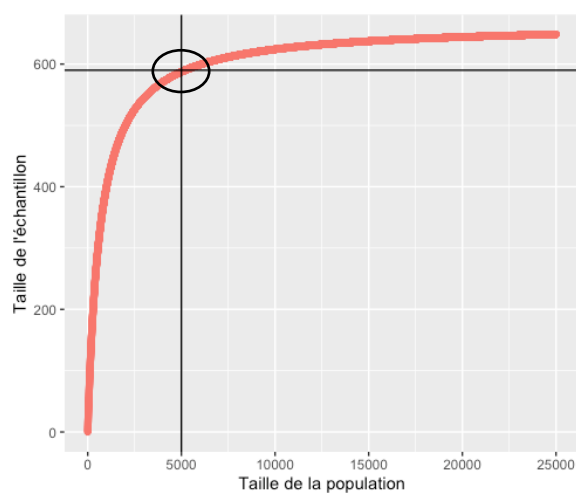


Figure 8. Taille d'échantillon pour un intervalle de confiance de 99 % et une marge d'erreur de 5 %

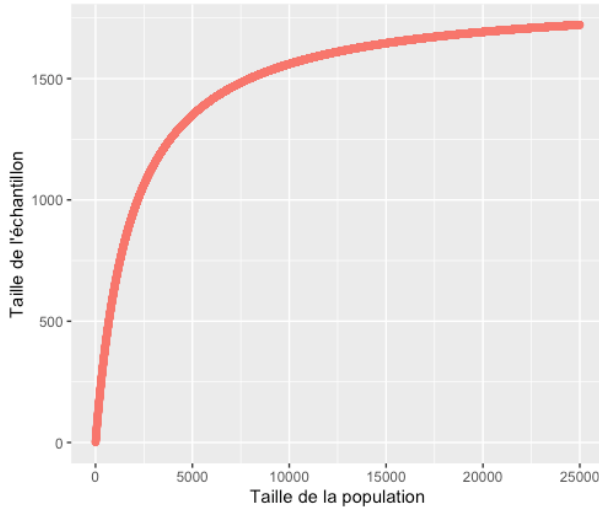


Figure 9. Taille d'échantillon pour un intervalle de confiance de 99 % et une marge d'erreur de 3 %

Les figures 10 à 12 illustrent la variation des tailles d'échantillon nécessaires selon la taille de la population pour un intervalle de confiance de 95 % et des marges d'erreur de 10 %, 5 % et 3 % respectivement. Pour une institution de sécurité comptant 5000 employé·e·s, par exemple, si l'on opte pour un niveau de confiance de 95 % et une marge d'erreur de 5 % afin d'évaluer l'institution, il faudra interroger environ 400 personnes.

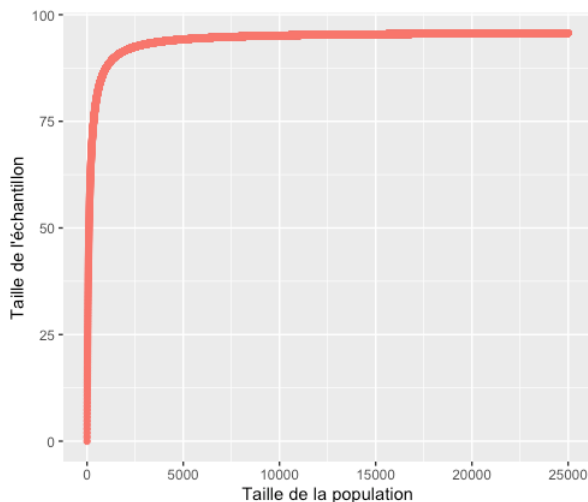


Figure 10. Taille d'échantillon pour un intervalle de confiance de 95 % et une marge d'erreur de 10 %

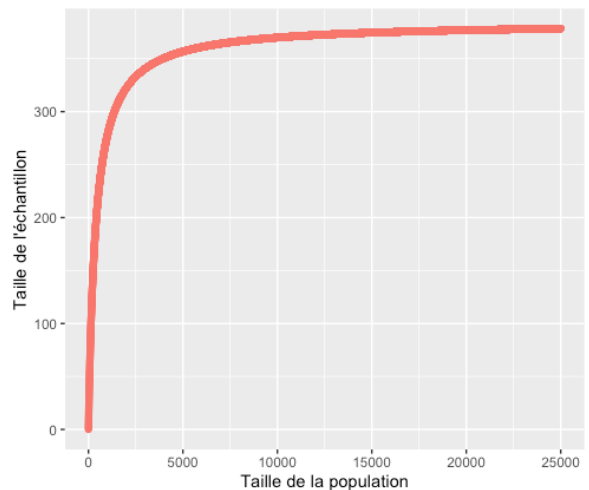


Figure 11. Taille d'échantillon pour un intervalle de confiance de 95 % et une marge d'erreur de 5 %

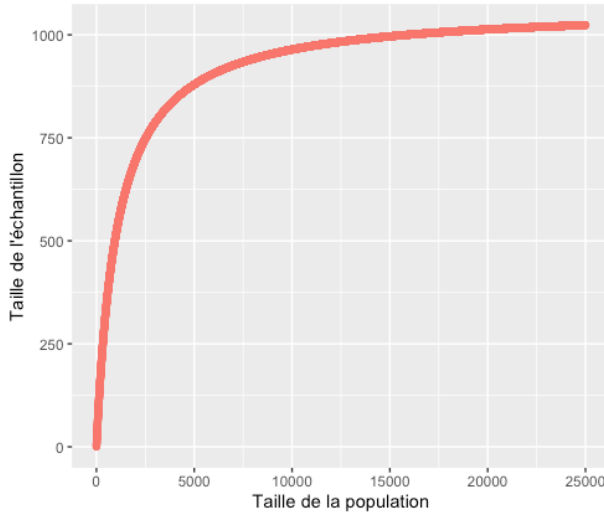


Figure 12. Taille d'échantillon pour un intervalle de confiance de 95 % et une marge d'erreur de 3 %

Les figures 13 à 15 montrent la variation des tailles d'échantillon nécessaires selon la taille de la population pour un intervalle de confiance de 90 % et des marges d'erreur de 10 %, 5 % et 3 % respectivement. Pour une institution de sécurité comptant 5000 employé·e·s, par exemple, si l'on opte pour un niveau de confiance de 90 % et une marge d'erreur de 5 % afin d'évaluer l'institution, il faudra interroger environ 300 personnes.

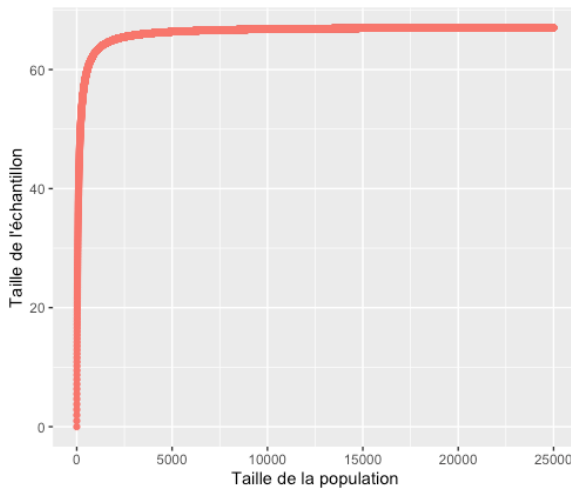


Figure 13. Taille d'échantillon pour un intervalle de confiance de 90 % et une marge d'erreur de 10 %

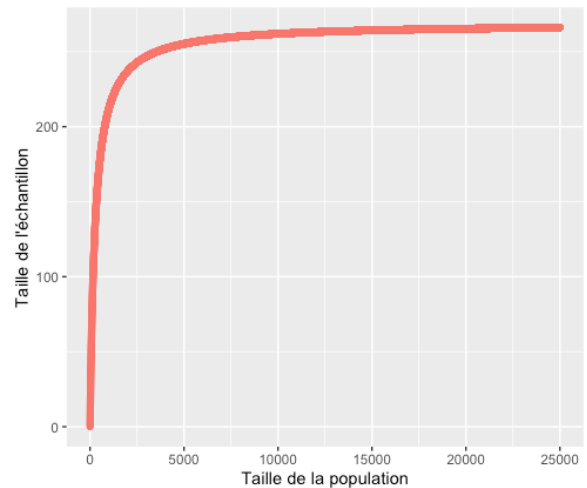


Figure 14. Taille d'échantillon pour un intervalle de confiance de 90 % et une marge d'erreur de 5 %

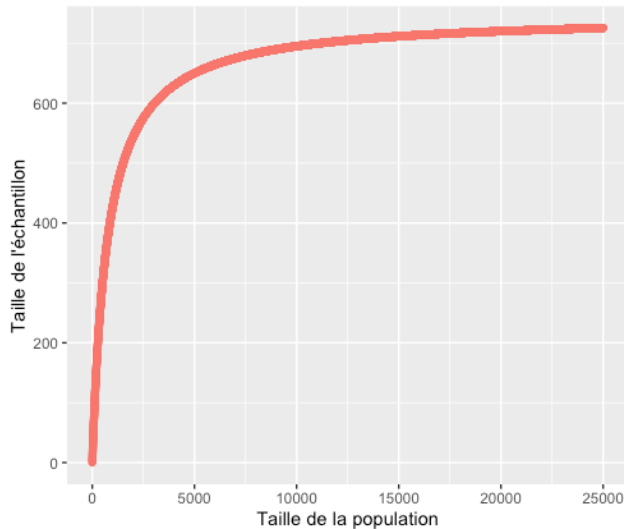


Figure 15. Taille d'échantillon pour un intervalle de confiance de 90 % et une marge d'erreur de 3 %

Section 4 : Comment créer une base d'échantillonnage

Cette section explique comment créer un échantillon de répondant·e·s dans le cadre de la méthodologie MOWIP.

Pourquoi utiliser une base d'échantillonnage ?

Pour obtenir une description plus précise des institutions de sécurité étudiées, il faut disposer d'un nombre suffisant de répondant·e·s provenant de différents groupes au sein de ces institutions. Ces personnes peuvent appartenir aux catégories suivantes :

- Personnel féminin et masculin
- Personnel de tous grades/rang élevé, moyen et inférieur
- Personnel déployé dans le cadre de missions de l'ONU et personnel non déployé
- Membres de l'armée de terre, de la marine et de l'armée de l'air (pour les institutions des forces armées)
- Personnel réparti sur plusieurs sites géographiques (provinces/États/régions d'un même pays)

Différentes méthodes de création d'une base d'échantillonnage

Il est possible de générer un échantillon à partir d'une liste principale (en l'occurrence, la base de données de l'institution contenant tous les membres du personnel en service) de différentes façons : en pratiquant une **sélection aléatoire**, un **échantillonnage systématique**, un **échantillonnage stratifié** ou un **échantillonnage par quotas**.

- **Sélection aléatoire à partir de la liste principale** : chaque membre de l'institution de sécurité se voit attribuer un numéro, et l'équipe d'évaluation ou les forces de sécurité utilisent un générateur de nombres aléatoires pour produire les numéros correspondant aux personnes qui seront interrogées.
- **Échantillonnage systématique à partir de la liste principale** : tous les membres de l'institution de sécurité sont répertoriés dans l'ordre alphabétique, soit par leur prénom soit par leur nom de famille (ou un autre critère tel que leur date de naissance), et l'équipe choisit un individu à intervalles réguliers, par exemple tous les 20 individus de la liste.
- **Échantillonnage stratifié à partir de la liste principale** : cette méthode permet de s'assurer que les caractéristiques nécessaires à l'enquête reflètent fidèlement la population de l'institution de sécurité. Si 30 % des forces de sécurité sont des femmes, par exemple, on veille à ce que l'échantillon compte 30 % de femmes.
- **Échantillonnage par quotas** : dans cette méthode, les groupes sélectionnés ne sont pas proportionnels à la population du groupe au sein de l'institution de sécurité. Si l'on s'intéresse à certaines caractéristiques, l'échantillon par quotas permet de garantir la surreprésentation des personnes présentant ces caractéristiques.

Pourquoi la méthodologie MOWIP recourt-elle en général à l'échantillonnage par quotas ?

La méthodologie MOWIP s'intéresse surtout aux expériences de certaines catégories de personnel : les **femmes** et les personnes **déployées/non déployées**. Pour cette raison, et parce qu'il faut que l'on interroge un nombre suffisant de ces personnes pour obtenir des résultats concluants, les groupes concernés sont suréchantillonnés. C'est-à-dire que même si les femmes ne représentent que 5 % de la population de l'institution de sécurité, on veille à ce qu'elles constituent 50 % de l'échantillon.

Instructions techniques pour la création d'une base d'échantillonnage

La première étape de la création d'une base d'échantillonnage par quotas consiste à recueillir les informations suivantes sur l'institution de sécurité étudiée :

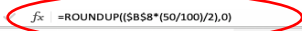
- Taille estimée de l'institution de sécurité
- Proportion (%) du personnel dans chaque division de l'institution
- Proportion (%) d'officiers·ères et de sous-officiers·ères dans chaque division (pour les corps d'armée uniquement)
- Localisation de tous les sites de l'institution (forces armées, police, gendarmerie) dans le pays et effectifs présents sur chaque site
 - Nombre d'hommes et de femmes sur chaque site
 - Effectifs déployés et non déployés sur chaque site
- Liste de toutes les unités non géographiques (unités opérationnelles, administratives, etc.)
 - Effectifs de ces unités
 - Nombre d'hommes et de femmes dans ces unités
 - Effectifs déployés et non déployés dans ces unités

Les informations obtenues sont ensuite importées dans une feuille de calcul Excel comme sur l'image ci-dessous. La méthodologie MOWIP recommande un échantillon minimum de 380 personnes pour chaque institution, ce qui signifie que l'échantillon comprend 95 femmes ayant été déployées, 95 femmes n'ayant pas été déployées, 95 hommes ayant été déployés et 95 hommes n'ayant pas été déployés. Les autres caractéristiques prises en compte dépendront de la proportion de personnel possédant cette caractéristique. Par exemple, dans le cadre d'une enquête sur les forces armées, on souhaite sélectionner des membres du personnel ayant le grade d'officier·ère et d'autres non. Le pourcentage d'officiers·ères et de sous-officiers·ères dans chaque division déterminera le nombre de personnes de ces deux catégories qui participeront à l'enquête. La figure ci-dessous présente un modèle simple de base d'échantillonnage pour les forces armées.

	A	B	C	D	E	F	G	H	I	
1	Fill yellow squares only.							Army	Navy	Air Force
2	1. What is the proportion (%) of personnel in each branch? (Change estimates).									
3						Total				
4						% Commissioned				
5						% NCO				
6										
7	Estimated size of military		(Excludes conscripts and paramilitary forces)							
8	Sample Size	380	(Will be higher as we round up.)							
9	Sample recently deployed	250								
10	Sample not recently deployed	250								
11										


Une fois les informations ci-dessus obtenues, un échantillon est créé pour les trois corps des forces armées. Comme on l'a vu, la méthodologie MOWIP utilise l'échantillonnage par quotas, ce qui implique le suréchantillonnage des femmes et du statut de déploiement. Les quotas liés au genre et au statut de déploiement sont déterminés comme suit : 50 % de femmes et 50 % d'hommes, et 50 % de personnel déployé et 50 % de personnel non déployé. Autrement dit, l'échantillon comprendra autant de femmes que d'hommes et autant de membres du personnel déployés que non déployés. L'échantillonnage par quotas permettra d'établir une comparaison significative et de la généraliser aux autres membres du personnel.

La figure ci-dessous montre comment constituer une base d'échantillonnage pour l'armée de terre. Comme ce corps d'armée rassemble la moitié du personnel des forces armées dans notre exemple, 50 % de l'échantillon lui est attribué. Autrement dit, 190 répondant·e·s viendront de l'armée de terre. Sur ces 190 personnes, 50 % seront des femmes et 50 % des hommes. Cela signifie qu'il y aura 95 femmes et 95 hommes de l'armée de terre participant à l'étude.

B17 

	A	B	C	D	E	F	G	H	I
1	Fill yellow squares only.						Army	Navy	Air Force
2	1. What is the proportion (%) of personnel in each branch? (Change estimates).								
3						Total			
4						% Commissioned			
5						% NCO			
6									
7	Estimated size of military		(Excludes conscripts and paramilitary forces)						
8	Sample Size	380	(Will be higher as we round up.)						
9	Sample recently deployed	190							
10	Sample not recently deployed	190							
11									
12									
13									
14									
15									
16						Total			
17						Women	Men		
18	Army	95		95					
19	Commissioned- been deployed	0		0					
20	Commissioned - never deployed	0		0					
21	NCO- been deployed	0		0					
22	NCO- never deployed	0		0					

Ces 95 femmes et 95 hommes seront ensuite réparti·e·s en quatre groupes : les personnes qui ont été déployées dans des missions de l'ONU et celles qui ne l'ont pas été, et les officiers·ères et les sous-officiers·ères. Le statut de déploiement ayant un quota de 50 %, ces 95 répondant·e·s seront divisé·e·s en deux groupes pondérés par le pourcentage total d'officiers·ères et de sous-officiers·ères dans le pays, comme le montre la figure ci-dessous. Dans cet exemple, il est supposé que 30 % du corps d'armée a le grade d'officier·ère et que 70 % ne l'a pas. Après avoir divisé 95 par 2, il faut pondérer le résultat par le nombre d'officiers·ères et de sous-officiers·ères pour obtenir l'échantillon final de personnes déployées et non déployées, et d'officiers·ères et de sous-officiers·ères dans l'armée de terre.

C18 

	A	B	C	D	E	F	G	H	I
1	Fill yellow squares only.						Army	Navy	Air Force
2	1. What is the proportion (%) of personnel in each branch? (Change estimates).								
3						Total			
4						% Commissioned	0.30		
5						% NCO	0.70		
6									
7	Estimated size of military		(Excludes conscripts and paramilitary forces)						
8	Sample Size	380	(Will be higher as we round up.)						
9	Sample recently deployed	190							
10	Sample not recently deployed	190							
11									
12									
13									
14									
15									
16						Total			
17						Women	Men		
18	Army	95		95					
19	Commissioned- been deployed	15		15					
20	Commissioned - never deployed	15		15					
21	NCO- been deployed	34		34					
22	NCO- never deployed	34		34					

Le même processus sera appliqué pour choisir l'échantillon total pour la marine et l'armée de l'air. Comme ces deux corps regroupent chacun environ 25 % du personnel dans notre exemple, 50 % de l'échantillon leur est attribué (25 % pour chaque corps). Autrement dit, 96 personnes viendront de la marine et 96 de l'armée de l'air. (On arrondit de 95 à 96 pour pouvoir diviser l'échantillon par deux.) Sur ces 96 personnes, 50 % seront des femmes et 50 % des hommes. Cela signifie qu'il y aura 48 femmes et 48 hommes de chaque corps participant à l'étude. Ces 48 femmes et

C60 $\text{fx} = \text{ROUNDUP}(\text{C59} * (\text{G54}) * 7/70, 0)$

	A	B	C	D	E	F	G	H	I
1	Fill yellow squares only.						Army	Navy	Air Force
2	1. What is the proportion (%) of personnel in each branch? (Change estimates).								
3						Total			
4						% Commissioned	0.30		
5						% NCO	0.70		
7	Estimated size of military		(Excludes conscripts and paramilitary forces)						
8	Sample Size	380	(Will be higher as we round up.)						
9	Sample recently deployed	190							
10	Sample not recently deployed	190							
58		Women	Men						
59	Army	95	95						
60	Commissioned- been deployed-Location 1	3	3						
61	Commissioned- been deployed-Location 2	3	3						
62	Commissioned- been deployed-Location 3	3	3						
63	Commissioned- been deployed-Location 4	3	3						
64	Commissioned- been deployed-Location 5	3	3						
65	Commissioned- been deployed-Location 6	3	3						
66	Commissioned- been deployed-Location 7	3	3						
67	Commissioned- never deployed-Location 1	2	2						
68	Commissioned- never deployed-Location 2	2	2						
69	Commissioned- never deployed-Location 3	2	2						
70	Commissioned- never deployed-Location 4	2	2						
71	Commissioned- never deployed-Location 5	2	2						
72	Commissioned- never deployed-Location 6	2	2						
73	Commissioned- never deployed-Location 7	2	2						
74	NCO- been deployed-Location 1	7	7						
75	NCO- been deployed-Location 2	7	7						
76	NCO- been deployed-Location 3	7	7						
77	NCO- been deployed-Location 4	7	7						
78	NCO- been deployed-Location 5	7	7						
79	NCO- been deployed-Location 6	7	7						
80	NCO- been deployed-Location 7	7	7						
81	NCO-never deployed-Location 1	3	3						
82	NCO- never deployed-Location 2	3	3						
83	NCO- never deployed-Location 3	3	3						

Section 5 : Résolution de différents problèmes d'échantillonnage

Cette section explique comment appliquer la méthodologie MOWIP de façon concluante lorsque l'échantillonnage pose des difficultés.

Comment gérer les cas où le nombre de femmes est limité dans le secteur de la sécurité ? Comment gérer les cas où le nombre de femmes déployées est limité ?

Afin d'étudier de façon concluante une population et ses sous-ensembles, il faut disposer d'un nombre suffisant d'individus dans les sous-ensembles de population de l'échantillon. Dans le cas de la méthodologie MOWIP, ce sont les expériences du personnel féminin que l'on veut étudier. Pour pouvoir étudier ce sous-groupe, il faut avoir un groupe de comparaison, c'est-à-dire du personnel masculin. Nous voulons également étudier les obstacles au déploiement auxquels le personnel est confronté. Cela entraîne la création de deux autres sous-groupes : le personnel déployé et le personnel non déployé. Nous avons besoin de ces quatre groupes, car, pour présenter des conclusions valables sur les expériences des femmes, il faut comparer les résultats obtenus pour les femmes à ceux obtenus pour les hommes. De même, pour présenter des conclusions valables sur le personnel déployé ou non déployé, il faut avoir des réponses des deux sous-groupes. Il faut donc interroger un nombre suffisant de femmes et d'hommes, et de membres du personnel déployés et non déployés, pour pouvoir mener l'étude à bien.

La méthodologie MOWIP recommande d'inclure dans l'échantillon au moins 190 femmes et 190 personnes déployées. Cependant, il peut arriver que certains pays n'atteignent pas le nombre requis d'individus dans un sous-ensemble particulier (par exemple, si l'institution compte moins de 95 femmes ayant déjà été déployées qui sont actuellement en poste). Dans ce cas, la méthodologie MOWIP recommande d'interroger chaque individu de ce groupe afin de recueillir les expériences de la totalité du sous-ensemble. Autrement dit, si le nombre total de femmes dans l'organisation est inférieur à 190, la méthodologie recommande d'inclure TOUTES les femmes dans l'échantillon.

Il en va de même lorsque le nombre total de personnes ayant été déployées qui sont actuellement en poste dans l'organisation est inférieur à 190. Dans ce cas, la méthodologie MOWIP recommande d'inclure TOUTES les personnes déployées dans l'échantillon.

Exemple

L'institution de sécurité compte 5 000 employé·e·s, mais seulement 100 femmes. Seules 110 personnes ont été déployées (100 hommes et 10 femmes). Il est alors recommandé d'interroger les 100 femmes et les 110 personnes déployées au sein de la force de sécurité, tout en enquêtant auprès de 380 personnes au total :

	Femmes	Hommes
Déployé·e·s	10	100
Non déployé·e·s	90	180
Total	100	280