

Simplifions la statistique

Par Louis Blais, stat., ASSQ

DE QUOI LA TAILLE D'UN ÉCHANTILLON DÉPEND-ELLE?

La taille d'un échantillon dépend :

- de la précision recherchée

Pour atteindre une précision de 90 %, il faut parfois plusieurs échantillons et le prix à payer peut être élevé.

- du niveau de confiance

C'est l'élément de risque.

- de la variabilité ou de l'étendue des valeurs possibles

Plus l'étendue des valeurs possibles est grande, plus il faut d'unités d'échantillonnage.

- de la superficie totale de la plantation

La taille d'un échantillon n'est influencée par la superficie totale de la plantation que si celle-ci est petite. Cette influence se traduit par une diminution du nombre d'unités d'échantillonnage. Dans les autres cas, la superficie totale n'a aucune influence sur la taille de l'échantillon. C'est ce que nous allons démontrer dans le présent document.

Dans une plantation, l'unité d'échantillonnage qui est utilisée pour estimer la densité de reboisement est une placette à rayon fixe dans laquelle les plants sont dénombrés. Le nombre d'unités d'échantillonnage constitue la taille d'échantillon. Si le nombre de plants à l'intérieur de chaque placette était toujours le même, peu importe la superficie de la plantation, une seule placette serait nécessaire pour connaître la densité

de plants par hectare (plants/ha). Il n'y aurait alors aucune variabilité. Ce simple exemple démontre que la superficie totale de la plantation n'a pas réellement d'effet. Pour illustrer concrètement ce propos, utilisons le cas réel d'une plantation qui, en principe, a moins de variabilité qu'une forêt naturelle, et trouvons la taille d'échantillon pour estimer la densité de reboisement sur une base de 2 500 plants/ha.

DÉMONSTRATION

Des données provenant de deux plantations de 1,2 ha, dans lesquelles chaque plant a été localisé précisément, sont utilisées pour cet exercice. Les deux plantations ont été réunies, puisque les densités de chacune sont semblables (1 995 et 2 036 plants/ha). Au total, la plantation compte donc 2 016 plants/ha sur une superficie de 2,4 ha (240 m sur 100 m). Puisque la population est connue, l'utilisation de la moyenne et de la variance de celle-ci peut servir pour estimer la taille d'échantillon nécessaire. La variance est une mesure de la variabilité de la population.

On calcule la moyenne de la population en utilisant le nombre total de plants, divisé par la superficie du territoire. La détermination de la variance de la population est plus complexe. Le tableau 1 donne un aperçu de la variabilité du nombre de plants/ha selon six tailles de placettes. Les tailles de placettes testées sont de 4, 5, 10, 25, 50, 100 et 200 m². Ce tableau présente une étendue couvrant 95 % des valeurs de plants/ha qu'il est possible d'obtenir avec chaque surface de placette.

TABLEAU 1 Aperçu de la variabilité du nombre de plants/ha de chaque placette dans une plantation réelle

Surface de la placette (m ²)	Rayon de la placette (m)	Étendue des densités couvrant 95 % des valeurs possibles (plants/ha)
4	1,13	De 0 à 5 000
5	1,26	De 0 à 4 000
10	1,78	De 0 à 4 000
25	2,82	De 800 à 3 200
50	3,99	De 1 000 à 2 800
100	5,64	De 1 300 à 2 700
200	7,98	De 1 500 à 2 500

OBJECTIFS DU SONDAGE

Le tableau 1 montre que plus la superficie de la placette augmente, plus l'étendue des valeurs possibles diminue. L'étendue passe de 0 à 5 000 plants/ha pour une placette de 4 m², à une étendue de 1 500 à 2 500 plants/ha pour une placette de 200 m². À l'extrême limite, si le territoire était échantillonné au complet, il n'y aurait alors qu'une valeur et donc aucune variabilité.

Les tableaux 2 et 3 présentent le nombre de placettes nécessaire pour obtenir une précision de 85 à 90 % à des niveaux de confiance respectivement de 95 et 90 %. On y constate que la taille d'échantillon nécessaire diminue à mesure que la surface de la placette augmente.

TABLEAU 2 Nombre de placettes nécessaire pour obtenir un niveau de confiance de 95 % selon la précision désirée et la surface de la placette, dans une plantation réelle de 2,4 ha ayant une densité de 2 015 plants/ha

Surface de la placette (m ²)	Précision					
	85 %	86 %	87 %	88 %	89 %	90 %
4	111	127	148	173	206	249
5	89	103	119	140	166	201
10	45	52	60	70	84	101
25	18	20	24	28	33	40
50	10	11	13	15	18	21
100	5	6	7	8	10	12
200	3	3	4	5	5	6

TABLEAU 3 Nombre de placettes nécessaire pour obtenir un niveau de confiance de 90 % selon la précision désirée et la surface de la placette, dans une plantation réelle de 2,4 ha ayant une densité de 2 015 plants/ha

Surface de la placette (m ²)	Précision					
	85 %	86 %	87 %	88 %	89 %	90 %
4	78	90	104	122	145	176
5	63	72	84	98	117	142
10	32	37	42	50	59	71
25	13	15	17	20	23	28
50	7	8	9	11	13	15
100	4	5	5	6	7	8
200	2	3	3	3	4	5

De plus, pour passer d'un taux de précision de 85 % à un taux de 90 % (soit 5 % de précision supplémentaire), le nombre de placettes nécessaire doit plus que doubler. Autre constatation : il faut 12 placettes de 100 m² sur cette petite surface pour obtenir une précision de 90 % avec un niveau de confiance de 95 %. La raison pour laquelle il faut autant de placettes est directement liée à la variabilité sur le terrain.

Les chiffres du tableau 1 démontrent la variabilité sur le terrain. Par exemple, pour les placettes de 100 m², 95 % des valeurs possibles se situent entre 1 300 et 2 700 plants/ha, ce qui représente une bonne variabilité.

Cette variabilité est principalement due au fait que tout le territoire ne peut « accueillir » tous les plants prévus. En effet, certaines zones dont les conditions de sol sont mauvaises pour les plants, côtoient d'autres zones dont les conditions sont adéquates. Cette variabilité explique pourquoi il faut autant de placettes sur un si petit territoire. Si la variance est la même pour une surface de plantation plus grande, la taille d'échantillon nécessaire sera également la même, puisqu'elle dépend de la variance.

La figure ci-dessous permet de faire une comparaison à l'aide de données après reboisement (sur une base de 2 500 plants/ha) fournies au ministère des Forêts, de la Faune et des Parcs par les bénéficiaires de contrat d'approvisionnement et d'aménagement forestier.



Une des façons de réduire la taille d'échantillon est d'utiliser un niveau de confiance moindre (par exemple 90 %) ou de diminuer la précision. Cette deuxième option n'est pas recommandée parce qu'il serait plus difficile de prouver que l'entrepreneur est fautif en cas de litige. Par contre, on peut diminuer le niveau de confiance de l'échantillonnage et en augmenter le niveau de risque.

Pour conclure, le nombre d'unités d'échantillonnage dépend principalement de la variabilité observée sur le terrain et non de la superficie du territoire. En foresterie, cela est vrai dans tous les cas.

POUR EN SAVOIR PLUS...

Ministère des Forêts, de la Faune et des Parcs
Direction de l'aménagement et de l'environnement forestiers
5700, 4e Avenue Ouest
Québec (Québec) G1H 6R1
daef@mffp.gouv.qc.ca