Résumé pédagogique : Introduction
Rappels statistiques et géostatistique

Rédigé par Dany Lauzon, Ph.D — Polytechnique Montréal

Objectifs d'apprentissage

- Pouvoir expliquer l'utilité de la géostatistique dans le domaine des géosciences et spécialement en mine ;
- Introduire les notions fondamentales de probabilités et statistiques ;
- Comprendre les principes de base des modèles spatiaux et leur application à l'estimation et la simulation des variables géoscientifiques ;

Table des matières

1 Introduction

La géostatistique est une branche des statistiques qui s'intéresse à l'analyse et à la modélisation des phénomènes spatiaux (et parfois spatio-temporels), en tenant compte de la localisation des données. Elle permet de décrire, d'interpoler et de prédire des variables régionalisées, telles que la teneur d'un minerai dans un gisement [1]. Grâce à des outils mathématiques et probabilistes, elle offre des estimations de valeurs inconnues ainsi que des mesures d'incertitude, essentielles pour la prise de décisions éclairées. Initialement centrée sur des approches univariées, la géostatistique s'est enrichie de méthodes multivariées intégrant des données secondaires, afin d'améliorer la précision des modèles. Elle joue un rôle clé dans divers domaines, notamment pour optimiser les opérations minières, tout en prenant en compte les dimensions économiques, environnementales et sociétales.

La géostatistique repose sur deux grands principes : l'effet de support et l'effet d'information. Dans les années 1950, un ingénieur d'Afrique du Sud, Daniel Krige, à observer deux grand phénomènes qui lui ont permis de postuler deux questions :

— Pourquoi récupère-t-on toujours moins de métal lorsque l'on exploite des grands volume qu'avec des petits volumes (effet support) ?

— Pourquoi récupère-t-on toujours moins de ressources (i.e., métal) avec des estimations qu'avec les valeur réelles du gisement connue après sont exploitation (effet information) ?

La discipline a été développer autour de ces deux grandes questions et le krigeage, développé en 1960 par George Matheron, fut nommé en l'honneur de Daniel Krige, le père fondateur de la géostatistique.

1.1 Effet de support

Un paradigme important doit être pris en considération lors de la réalisation d'estimations et de leur utilisation dans les opérations minières courantes. Les forages exploratoires et d'exploitation ont des diamètres très petits (quelques centimètres) par rapport à la taille des blocs (quelques mètres) qui sont exploité par la minière. On exploite donc une ressource sur des blocs, mais leurs estimations reposent sur des carottes. On appel la taille étudiée est le support. Ainsi, le support de nos observations est celui des forages, soit un petit support, tandis que le support de nos estimations doit être celui du bloc, soit un grand support. Il est donc crucial de prendre en compte le support lors des estimations et procéder correctement au changement de support. Cela sera traité dans un autre leçon.

La Fig.1 présente deux gisements miniers ayant une distribution statistique des teneurs des forages (petit support) identique, c'est-à-dire que leur histogramme est identique (ou que leur fonction de densité ou de répartition est similaire). Par conséquent, le calcul des ressources est identique lorsque l'on considère le forage comme support. Cependant, il est impossible d'opérer sur un support de quelques centimètres ; il suffit de regarder la taille des équipements miniers pour comprendre que

[1] La géostatistique permet également de modéliser et d'analyser des séries temporelles. Dans ce contexte, la variable d'intérêt n'est pas nécessairement la teneur en fonction de la localisation dans un gisement, mais peut être la teneur du minerai envoyé au concentrateur pour y être traité. On s'intéresse alors à la stabilité de cette teneur au cours du temps, c'est-à-dire à sa variabilité éventuelle, dans le but d'optimiser les procédés chimiques utilisés pour extraire le métal de la roche encaissante.

l'on opère sur des blocs de taille de l'ordre des mètres. Ainsi, lorsque l'on augmente la taille du support, en passant du forage à un bloc de taille plus importante, on modifie les statistiques de nos teneurs, et la nature du gisement influencera ces statistiques. On constate que l'histogramme des teneurs des blocs des gisements A et B est complètement différent lorsque l'on augmente la taille du bloc. Cela montre que la teneur des carottes de forage ne permet pas d'expliquer toute la complexité du gisement et que d'autres phénomènes interviennent, tels que la continuité spatiale du gisement et la structure minéralogique.

En règle générale, on récupère toujours moins de métal avec de gros blocs qu'avec de petits blocs. Ce phénomène est connu sous le nom d'effet de support. Pourquoi ? Parce que la minéralisation est un phénomène sporadique. Plus on augmente la taille du bloc, plus on introduit des concentrations faibles, car on a moins de chances d'avoir plusieurs zones riches dans un même bloc. Par conséquent, en raison de l'effet de moyenne, la teneur des grands blocs sera toujours inférieure à celle des blocs plus petits, bien entendu en moyenne. Il est possible que la teneur d'un bloc augmente, mais en moyenne, celle-ci tend à diminuer. Ce phénomène est observable dans les histogrammes lorsque l'on compare les différentes tailles de blocs, de haut en bas.

1.2 Effet d'information

La quantité de forages joue également un rôle clé dans l'estimation des ressources. Il est évident qu'il est beaucoup plus facile d'estimer une valeur à une position donnée avec un million d'observations qu'avec une seule. En effet, plus nous disposons de données, plus nos estimations seront fidèles à la réalité se trouvant sous nos pieds. Ce phénomène est connu sous le nom d'effet d'information.

Cependant, cet effet n'est pas entièrement lié à la quantité de données. Il est tout aussi important de considérer leur qualité. Dans de futures lectures, nous démontrerons qu'il est crucial de bien positionner les forages afin de maximiser le gain d'information tout en limitant le nombre de forages. En somme, ajouter un forage supplémentaire dans une petite zone déjà densément couverte n'apporte pas nécessairement beaucoup d'information additionnelle. Le gain serait faible, et il conviendrait alors de réfléchir (voire de calculer) son positionnement pour maximiser son utilité ailleurs dans la mine.

Par ailleurs, les informations issues des forages sont souvent entachées d'erreurs : erreurs d'analyse des teneurs, erreurs de localisation des carottes dans l'espace, erreurs de modélisation numérique, etc. En fin de compte, nos observations ne correspondent pas toujours à la teneur réelle sur le terrain. Il y aura donc toujours une forme de biais ou d'erreur, ce qui empêche de garantir que la teneur mesurée est égale à la teneur réelle.

Ainsi, on récupère toujours moins de métal avec des estimations qu'avec les vraies valeurs, car les décisions sont prises à partir d'estimations imparfaites, tandis que l'exploitation repose sur la réalité géologique. Ce principe peut être relié aux notions de faux positifs et faux négatifs.

La Figure 2 illustre de façon simplifiée les différences entre les teneurs estimées et les teneurs réelles, mesurées après exploitation. Nous prenons nos décisions en fonction des estimations : tout le matériel situé à droite de la ligne verticale sera donc traité. Mais, puisque nos estimations ne sont pas parfaites, une certaine quantité de stérile (section brune) sera également traitée, en raison des erreurs d'estimation.

Cette illustration ne tient pas compte des biais conditionnels et des biais systématiques présents dans les estimations. Nous aborderons ces notions plus en détail lors de l'étude des méthodes d'estimation. Il est cependant essentiel de garder en tête que nous récupérons toujours moins de minerai lorsque les décisions sont prises à partir d'estimations, et que celles-ci comportent inévitablement des incertitudes. Comme nos décisions sont toujours basées sur des estimations, il devient crucial d'utiliser des méthodes d'estimation rigoureuses, précises et sans biais.

1.3 L'utilité de la géostatistique

La géostatistique permet de prévoir l'ampleur des effets de support et d'information, afin de minimiser leur impact sur les opérations minières et de prendre des décisions éclairées. Grâce à des analyses spatiales des statistiques (moyenne, variance, covariance, corrélation, intervalle de confiance), il est possible de quantifier ces effets à l'aide de modèles mathématiques théoriques. La géostatistique combine ainsi deux branches des mathématiques : les probabilités et statistiques, et l'algèbre linéaire. En appliquant ces disciplines à la géologie et aux opérations minières, elle offre un cadre d'analyse puissant.

2 Rappel de probabilités et statistiques

La géostatistique est une branche de la statistique qui se concentre sur les ensembles de données spatiales ou spatio-temporelles. Elle permet, par exemple, de prédire les distributions de probabilité des teneurs en minerai dans le cadre des opérations minières. Il est donc important de revoir les bases en probabilité et en statistique afin de maîtriser la terminologie et les concepts mathématiques propres à cette discipline.

2.1 Variable aléatoire, fonction de densité et fonction de répartition

Une variable aléatoire (v.a.) est une fonction mathématique qui associe un résultat numérique à chaque issue possible d'une expérience aléatoire. Les valeurs possibles de la v.a. sont connues, mais le résultat précis d'une

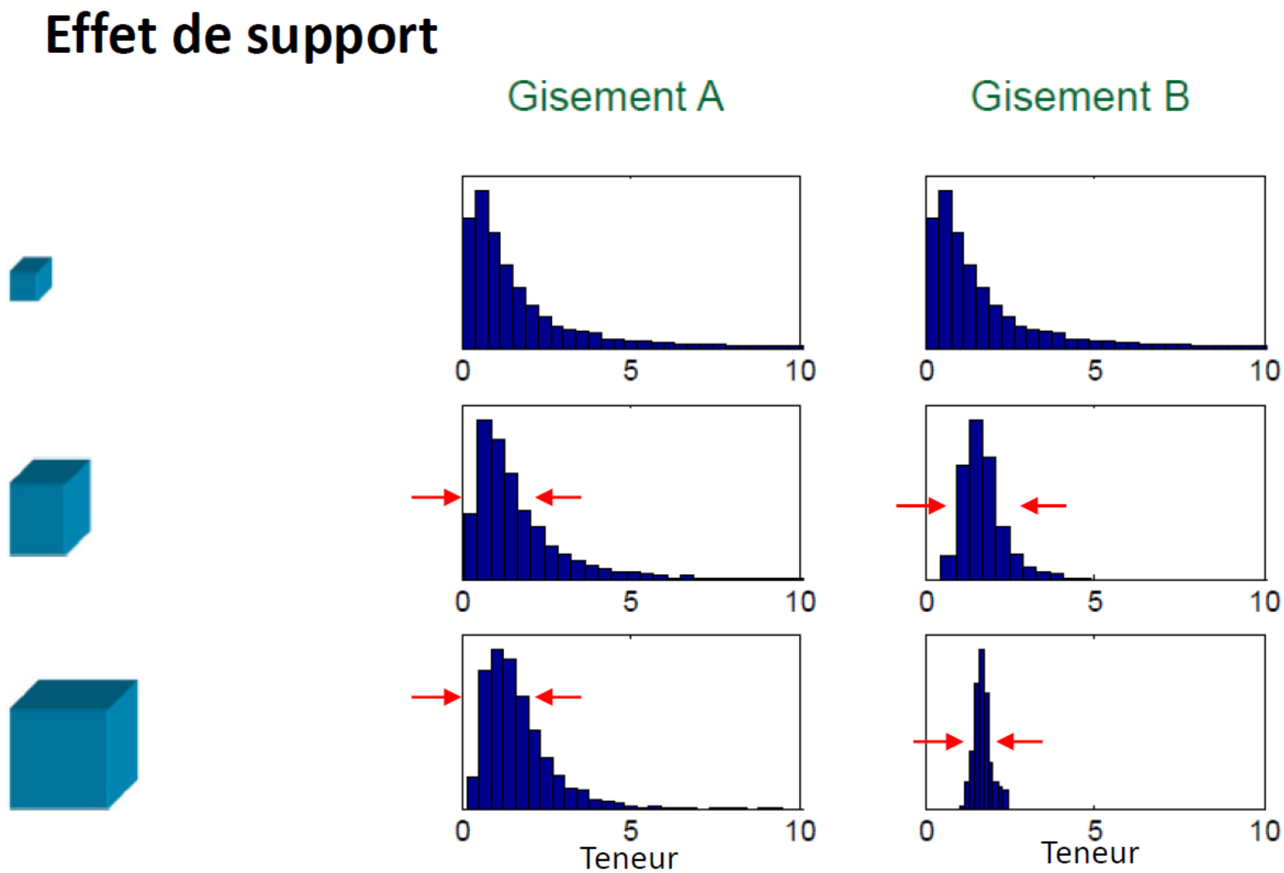

Figure 1 –

réalisation ne peut être déterminé *a priori* sans observation direct. Par exemple : la teneur en cuivre d'une carotte de forage de 1 mètre, l'épaisseur d'une veine minéralisée, la concentration d'un polluant dans une nappe phréatique ou encore le pH de l'eau de pluie.

Même si la valeur exacte que prendra une variable aléatoire n'est pas connue, il est possible d'estimer la probabilité qu'elle prenne certaines valeurs.

Cette information est décrite à l'aide :

— de la **fonction de masse** $p_X(x)$, pour les v.a. **discrètes**, et

— de la **fonction de densité** $f_X(x)$, pour les v.a. **continues**.

Dans le cadre de nos lectures, nous travaillerons uniquement avec des variables aléatoires continues.

La fonction de densité $f_X(x)$ vérifie deux propriétés essentielles :

1. Elle est **positive** partout :

$$f_X(x) \geq 0 \quad \text{pour tout } x \in \mathbb{R}$$

2. L'**aire sous la courbe** est égale à 1 (probabilité totale) :

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1$$

La probabilité que la variable aléatoire prenne une valeur comprise entre deux bornes a et b, soit $P(a \leq X \leq b)$, est donnée par l'intégrale de la fonction de densité entre ces deux bornes :

$$P(a \leq X \leq b) = \int_a^b f_X(x)\, dx$$

Cette relation conduit naturellement à la définition de la fonction de répartition, notée $F_X(x)$, qui donne la probabilité que la variable aléatoire X prenne une valeur inférieure ou égale à x :

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} f_X(t)\, dt$$

La fonction de répartition est une fonction croissante, bornée entre 0 et 1, et continue pour les variables continues. Elle est particulièrement utile pour visualiser la distribution cumulative des probabilités et pour déterminer des quantiles, comme la médiane (valeur pour laquelle $F_X(x) = 0{,}5$).

Mise en contexte

Soit $Z(x)$ une variable aléatoire représentant la valeur de la variable d'intérêt à une position x. Cette valeur est inconnue (par exemple : teneur, température,

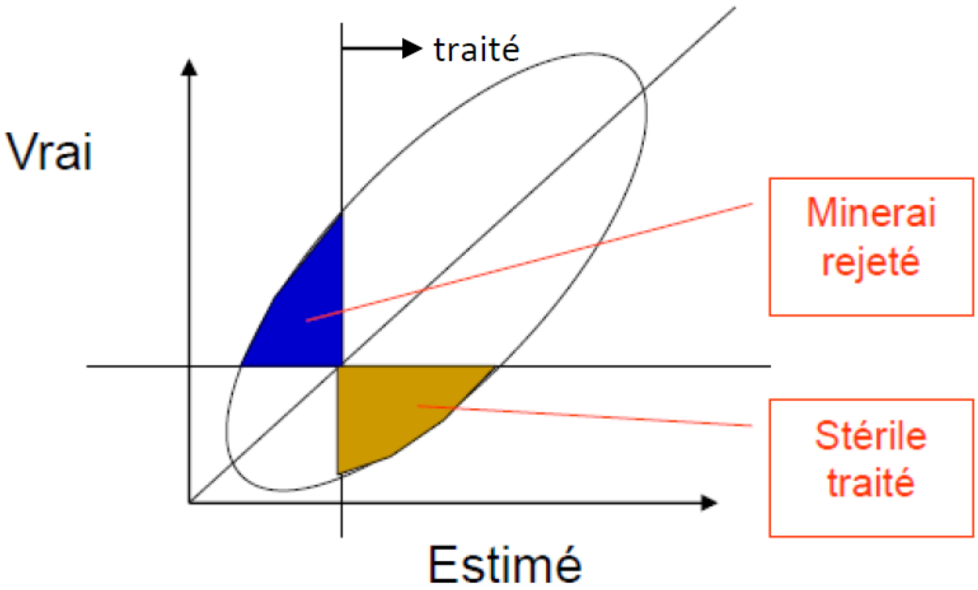

Figure 2 – Illustration des erreurs d'estimation : comparaison entre les teneurs estimées et les teneurs réelles.

précipitations, niveau piézométrique, etc.). Bien qu'il existe une valeur réelle en ce point x, que l'on pourrait mesurer, la géostatistique considère cette valeur comme aléatoire, car elle n'a pas été mesurée (ou ne l'a pas encore été). On définit donc, à partir de la fonction de répartition, les plages possibles que peut prendre cette valeur en fonction des informations disponibles concernant $Z(x)$:

$$F(z, x) = \text{Prob}\,\{Z(x) \leq z \mid \text{informations}\}$$

Il est important de noter que, dans cet exemple, la fonction de répartition dépend de la localisation x de la variable. On rappelle que la géostatistique s'intéresse à des variables aléatoires régionalisées, c'est-à-dire ayant une position spatiale ou temporelle.

2.2 Mesures de tendance centrale

Les mesures de tendance centrale permettent de résumer une distribution de probabilité par une valeur représentative de l'ensemble des résultats possibles. Voici les principales :

— **Mode** : valeur x pour laquelle la fonction de densité $f_X(x)$ est maximale. Il s'agit du point le plus probable :

$$\text{Mode} = \arg\max_x \, f_X(x)$$

— **Médiane** : valeur x telle que la moitié des observations se situent en dessous de cette valeur :

$$P(X < x) = 0{,}5 \quad \Leftrightarrow \quad \int_{-\infty}^{x} f_X(t)\, dt = 0{,}5$$

— **Moyenne (ou espérance)** : valeur moyenne attendue de la variable aléatoire, notée μ :

$$\mu = E[X] = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

Ces mesures peuvent différer selon la forme de la distribution. Par exemple, pour une distribution symétrique comme la loi normale, la moyenne, la médiane et le mode coïncident. Pour des distributions asymétriques (ex. : loi log-normale), ces mesures seront différentes.

2.3 Mesures de dispersion

Les mesures de dispersion sont des statistiques qui permettent de décrire la variabilité ou l'étendue des valeurs d'une variable aléatoire autour de sa moyenne. Ces mesures sont essentielles pour comprendre l'incertitude associée à la variable étudiée et pour évaluer à quel point les données sont concentrées ou dispersées.

— **Variance :** La variance mesure l'étendue des valeurs par rapport à la moyenne et est définie par :

$$\sigma_X^2 = E[(X - \mu)^2] = \int_{-\infty}^{\infty} (x - \mu)^2 f_X(x)\, dx$$

Elle donne une idée de l'écart moyen quadratique entre les valeurs observées et la moyenne.

— **Écart-type :** L'écart-type est la racine carrée de la variance et permet d'exprimer la dispersion des données dans les mêmes unités que la variable elle-même :

$$\sigma_X = \sqrt{\sigma_X^2}$$

C'est une mesure plus intuitive de la variabilité des données.

— **Asymétrie :** L'asymétrie mesure l'asymétrie de la distribution par rapport à la moyenne. Si l'asymétrie est positive, la distribution est étendue vers la droite, tandis que si elle est négative, elle est étendue vers la gauche. Elle est calculée par :

$$\frac{E[(X-\mu)^3]}{\sigma_X^3}$$

— **Aplatissement :** L'aplatissement indique le degré de "plateur" ou de "pic" de la distribution par rapport à une distribution normale. Une valeur élevée indique une distribution plus pointue, tandis qu'une valeur faible indique une distribution plus aplatie. Il est défini par :

$$\frac{E[(X-\mu)^4]}{\sigma_X^4}$$

2.4 Estimation à partir d'un échantillon

L'estimation à partir d'un échantillon consiste à utiliser les données observées pour estimer les caractéristiques d'une population ou d'une distribution inconnue. Les estimations reposent sur des mesures statistiques calculées à partir de l'échantillon disponible. Cela inclut des estimations des paramètres de tendance centrale, de dispersion, et de la distribution de la variable.

— **Moyenne empirique :** La moyenne empirique est l'estimation de la moyenne de la population à partir d'un échantillon. Elle est définie par :

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

C'est la somme des valeurs observées divisée par le nombre total d'observations. La moyenne empirique est utilisée pour estimer la tendance centrale de la variable aléatoire.

— **Variance empirique :** La variance empirique permet d'estimer la variabilité des données dans l'échantillon par rapport à la moyenne empirique. Elle est donnée par :

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

Cette estimation corrige le biais de sous-estimation de la variance de la population en utilisant $n-1$ au lieu de n.

— **Fonction de densité :** La fonction de densité empirique peut être estimée à l'aide d'un histogramme. Un histogramme est une représentation graphique de la distribution des données, divisée en intervalles. Il donne une idée de la forme de la distribution et de la densité des observations dans chaque intervalle.

— **Fonction de répartition empirique :** La fonction de répartition empirique, notée $F(x)$, donne la proportion d'observations inférieures ou égales à une valeur donnée x. Elle est calculée comme suit :

$$F(x) = \frac{\text{rang}(x_i)}{n}$$

où $\text{rang}(x_i)$ est le rang de x_i dans l'échantillon, c'est-à-dire la position de x_i lorsque les données sont triées dans l'ordre croissant. Cette fonction permet de visualiser la distribution cumulative des données.

— **Estimateur sans biais** : Un estimateur est dit sans biais si son espérance mathématique est égale à la valeur réelle du paramètre qu'il estime. En d'autres termes, un estimateur sans biais fournit, en moyenne, une estimation correcte du paramètre cible. Par exemple, l'estimateur de la teneur vraie $Z^{\text{vraie}}(x)$ est dit sans biais si, en moyenne, notre estimé $Z^*(x)$ retourne la valeur vraie :

$$E[Z^*(x)] = Z^{\text{vraie}}(x)$$

2.5 Fonction de densité conjointe

Lorsqu'on considère plusieurs variables aléatoires, comme dans le cas de deux variables X et Y, leur dépendance conjointe peut être représentée par une fonction de densité conjointe, notée $f_{XY}(x,y)$. Cette fonction décrit la probabilité que $X = x$ et $Y = y$ se produisent simultanément.

La condition de normalisation de cette fonction de densité conjointe est la suivante :

$$\iint f_{XY}(x,y)\,dx\,dy = 1$$

Dans le cas de deux variables aléatoires, les mesures usuelles de dépendance sont la covariance et la corrélation :

— Covariance :

$$\text{Cov}(X,Y) = E[(X-E[X])(Y-E[Y])]$$

— Corrélation :

$$\rho_{XY} = \frac{\text{Cov}(X,Y)}{\sigma_X\sigma_Y} \in [-1,1]$$

Il est à noter que la corrélation est une normalisation de la covariance, ce qui permet d'obtenir une plage de valeurs comprises entre $[-1,1]$. Une valeur de corrélation de 1 indique que les valeurs de X et Y sont entièrement dépendantes l'une de l'autre, c'est-à-dire qu'elles varient de manière parfaitement linéaire dans la même direction. Une valeur de 0 indique qu'il n'y a aucune dépendance linéaire entre X et Y, c'est-à-dire qu'ils sont indépendants. Enfin, une valeur de corrélation de -1 signifie que la relation entre X et Y est inverse, c'est-à-dire que lorsque X augmente, Y diminue de manière parfaitement

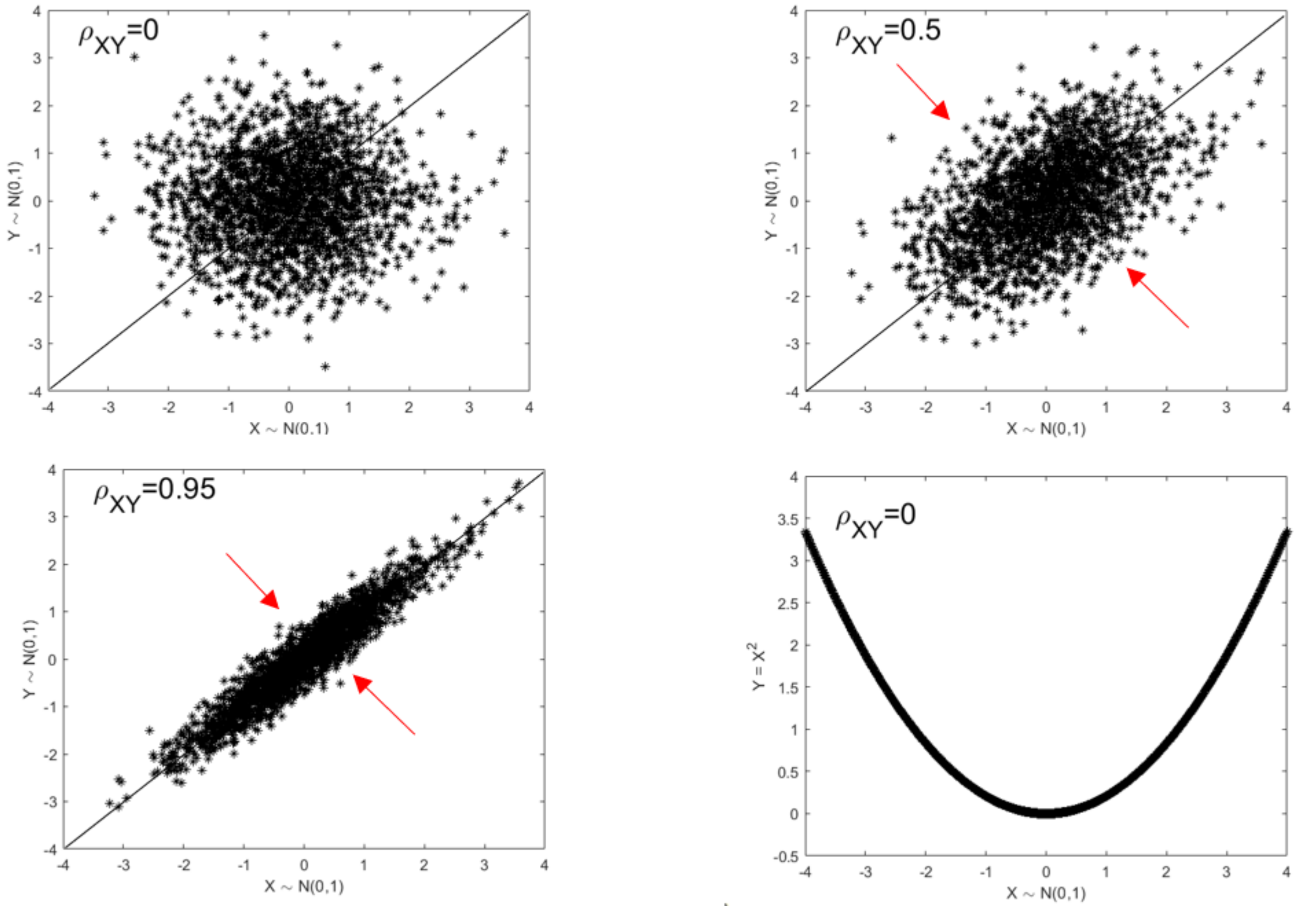

FIGURE 3 – Exemple de corrélation entre deux variables aléatoires X et Y.

linéaire. La Fig.3 montre différent scénario de corrélation entre deux variables.

Il est également important de noter que si X et Y sont indépendantes, alors la covariance entre X et Y est nulle, c'est-à-dire :

$$\mathrm{Cov}(X, Y) = 0$$

Propriétés

La variance de la somme de deux variables aléatoires X et Y est donnée par :

$$\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

De plus, la variance d'une combinaison linéaire de X et Y est donnée par :

$$\mathrm{Var}(aX + bY) = a^2\,\mathrm{Var}(X) + b^2\,\mathrm{Var}(Y) + 2ab\,\mathrm{Cov}(X, Y)$$

Enfin, la variance de la somme pondérée de n variables aléatoires $X_1, X_2, \ldots, X_n$ est donnée par :

$$\mathrm{Var}\left(\sum_{i=1}^{n} a_i X_i\right) = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i a_j\,\mathrm{Cov}(X_i, X_j)$$

Ces relations sont au cœur même de la géostatistique et constituent très probablement la relation la plus utilisée dans le domaine. Il est donc crucial de bien les maîtriser.

En effet, la compréhension de la covariance, de la corrélation et de leurs applications dans les modèles géostatistiques est essentielle pour analyser et interpréter correctement les données spatiales et spatio-temporelles. Ces concepts permettent de quantifier et de modéliser la dépendance entre différentes variables, ce qui est fondamental dans la prédiction des valeurs à des localisations non échantillonnées.

2.6 Interprétation géostatistique

Les v.a en géostatistique sont régionalisées, c'est-à-dire qu'elles dépendent de leur position spatiale, notée $Z(x)$, où chaque valeur est associée à une localisation particulière dans l'espace ou dans le temps. Cette dépendance spatiale est fondamentale en géostatistique, car elle permet de modéliser et de prédire des valeurs à des positions non mesurées en fonction des données disponibles à d'autres positions.

Supports

Le support fait référence à la zone ou à la région sur laquelle les mesures sont effectuées. En fonction de la taille et de la nature du support, les statistiques associées à la variable aléatoire peuvent varier.

— **Ponctuel :** Si la mesure est effectuée en un seul point, on a la valeur moyenne de la variable Z_G sur un support ponctuel, défini par :

$$Z_G = \frac{1}{|G|} \int_G Z(x)\, dx$$

où $|G|$ représente la taille du domaine G et $Z(x)$ est la valeur de la variable aléatoire à la position x. Le support ponctuel est généralement la taille de la carotte.

— **Petits blocs :** Si les mesures sont effectuées sur un petit bloc de données, la moyenne de $Z(x)$ est définie par :

$$\bar{Z}_G = \frac{1}{N} \sum_{i=1}^{N} Z(x_i)$$

où N est le nombre d'observations dans le petit bloc G, et x_i représente les positions de chaque observation.

— **Gros blocs :** Pour de plus grandes régions, la moyenne $\bar{Z}_G$ est définie par :

$$\bar{Z}_G = \frac{1}{M} \sum_{i=1}^{M} Z(x_i)$$

où M représente le nombre de mesures dans un gros bloc, et x_i les positions de ces mesures.

Les statistiques calculées sur la variable aléatoire $Z(x)$ dépendent directement du support de mesure. En particulier, la variance de la variable aléatoire a tendance à décroître à mesure que la taille du support augmente. En effet, plus le support est grand, plus la variabilité des valeurs mesurées est réduite, ce qui donne lieu à des estimations plus stables et moins sensibles aux fluctuations locales.

Cependant, la taille optimale du support dépend généralement des opérations minières et des équipements disponibles pour exploiter cette taille. Il n'est donc pas correct de dire que plus la taille du support augmente, plus les profits seront importants. La relation entre la taille du support et les profits est plus complexe et doit prendre en compte divers facteurs tels que l'efficacité des équipements, la nature géologique du gisement et les coûts associés à l'exploitation. Nous explorerons certaines de ces causes dans les prochaines lectures.

3 Applications de la géostatistique dans l'exploration et l'exploitation minière

Dans le secteur **minier**, la géostatistique intervient dès les premières étapes d'un projet, notamment pour estimer les ressources et évaluer la faisabilité économique. En phase d'exploitation, elle permet de guider quotidiennement le tri du minerai, en déterminant quels matériaux doivent être envoyés à l'usine de traitement et lesquels sont considérés comme des stériles, à partir des données les plus récentes. On résume certaine domaine d'application dans le secteur minier ci-dessous :

— **Planification minière, délimitation de chantiers** : Cette étape consiste à définir et organiser les zones d'extraction dans une mine, en fonction des ressources disponibles et des objectifs de production. Cela permet de déterminer les zones à exploiter et les priorités d'extraction.

— **Évaluation économique de scénarios d'exploitation** : Cette tâche implique l'analyse de différents scénarios d'exploitation minière pour évaluer leur rentabilité. Elle prend en compte les coûts de production, les revenus potentiels et les risques associés à chaque scénario.

— **Séquence d'exploitation optimale** : Il s'agit de déterminer la meilleure manière de planifier l'extraction des ressources, en tenant compte de facteurs comme la rentabilité, la gestion des risques et l'impact environnemental, pour maximiser l'efficacité de l'exploitation.

— **Détermination des contours d'une fosse optimale** : Cela consiste à définir les limites géométriques d'une fosse d'exploitation en fonction des ressources et des contraintes techniques et économiques, de manière à maximiser l'extraction tout en minimisant les coûts.

— **Analyse de variabilité pour concentrateurs** : Cette analyse évalue la variabilité des propriétés du minerai (comme la teneur en métaux) pour optimiser les processus de concentration et de traitement dans les concentrateurs, afin d'améliorer l'efficacité de l'extraction des métaux.

— **Homogénéisation du minerai extrait** : Cette étape vise à uniformiser la qualité du minerai extrait afin d'éviter les variations importantes dans les teneurs, ce qui permet une gestion plus stable des processus de traitement et d'optimisation de la production.

— **Prédictions à court terme de la teneur** : Il s'agit d'estimer la teneur du minerai extrait sur une période à court terme, afin d'ajuster les processus de traitement et de planifier la production de manière plus précise, ce qui aide à optimiser la gestion des ressources.

Récemment, on s'intéresse à l'utilisation de la géostatistique pour l'optimisation de la conception et de la construction des haldes à stériles et des parcs à résidus dans la restauration des sites miniers. Il s'agit d'un jeune domaine de recherche dont les applications sont très prometteuse. Il s'agit de travaux de recherche en cours à Polytechnique Montréal à travers l'Institut de la recherche en mines et environnement.

4 Autres domaines d'application

La géostatistique n'est pas limiter à des applications en mines. Elle est largement utilisée dans de nombreux domaines des sciences et du génie. Par exemple :

Dans le domaine de la **géotechnique**, la géostatistique est utilisée pour modéliser en 2D ou en 3D la géologie du sous-sol et estimer, par la suite, les propriétés associées à chacune des unités. Elle cherche à répondre à une double question : où se trouvent les unités argileuses ou sableuses, et comment estimer spatialement leurs propriétés mécaniques et hydrogéologiques ? On appelle cette approche la modélisation de faciès. Ainsi, la géostatistique permet d'obtenir un modèle du sous-sol cohérent avec les données disponibles, offrant ainsi une approche plus rigoureuse pour la conception des modèles géotechniques, qui étaient auparavant basés sur des méthodes empiriques et l'expérience de l'ingénieur. En utilisant un nombre suffisant de données, elle garantit la pertinence des analyses, même si, dans certains cas, les données disponibles peuvent être limitées [2].

La **géomécanique**, dont la modélisation des réseaux de fractures bénéficie largement de la géostatistique, utilise cette approche pour l'estimation et la modélisation de la densité des réseaux de fractures, en particulier pour des applications en stabilité des pentes et en écoulement des eaux dans ces fractures. L'une des applications consiste à décrire la distribution spatiale des fractures à partir de données de forages, d'imagerie ou de relevés géophysiques, en quantifiant leur densité, leur orientation, leur connectivité, leur ouverture et leur rugosité. Par la suite, la méthode de modélisation est sélectionnée en fonction des besoins et de l'étude géostatistique.

En **sciences de l'environnement**, elle est utilisée pour évaluer la concentration de polluants dans les sols, les eaux ou l'air, afin d'estimer les risques pour la santé humaine et l'écosystème, et de déterminer la nécessité d'interventions de dépollution.

Dans le domaine de la **science des sols**, des applications récentes s'intéressent à la cartographie de nutriments (azote, phosphore, potassium, etc.) et d'indicateurs comme la conductivité électrique, dans une optique d'agriculture de précision. Ces cartes permettent de moduler finement les apports en engrais selon les besoins spécifiques de chaque zone du champ.

Les **applications météorologiques** mobilisent également la géostatistique pour la prévision de variables telles que la température, les précipitations ou des phénomènes connexes comme les pluies acides, à partir de données d'observation ponctuelles [3]

Dans le domaine de la **santé publique**, la géostatistique est de plus en plus utilisée pour modéliser la répartition spatiale de contaminants environnementaux et leur lien avec des indicateurs sanitaires, comme les taux d'incidence de certaines maladies.

Pour continuer la liste, on note des applications dans de nombreux autres domaines tels que les gisements pétroliers, la résolution de problèmes inverses en géophysique, la cartographie assistée, la classification des sols, l'analyse d'images et bien d'autres domaines. L'important est d'avoir des coordonnées spatiales ou temporelles et des valeurs observées.

Dans tous ces contextes, l'objectif est de caractériser un phénomène dont la mesure exhaustive serait trop coûteuse ou chronophage, à l'aide de données collectées en un nombre restreint de points. Les méthodes géostatistiques permettent alors de produire des cartes de prédiction accompagnées de mesures d'incertitude, afin de mieux comprendre le phénomène et de soutenir les décisions opérationnelles, environnementales ou économiques.

2. L'un de mes projets en cours se concentre sur la modélisation complète des unités meubles du sous-sol québécois à travers une base de données contenant plus de 300 000 forages géotechniques.

3. Il s'agit d'un autre axe de mes travaux de recherche où je m'intéresse à modéliser les comportements asymétriques observés dans les champs de précipitations.