HEC MONTRÉAL

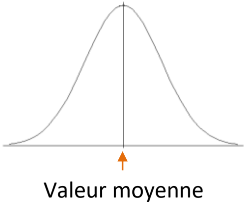
30-650-17
Introduction à l'analytique d'affaires

Thème 4 :
Outils d'aide à la décision en contexte d'incertitude

La loi normale

Sonnez les cloches: voici la loi normale!

La loi normale est une distribution en forme de <u>cloche symétrique</u>. Elle est souvent appropriée pour décrire le comportement de variables dont la plupart des observations se situent autour de leur moyenne, en proportions égales au-dessus et en-dessous de la valeur moyenne.



Exemples: Voici quelques variables dont le comportement peut adéquatement être décrit par la loi normale :

- plusieurs variables continues relatives à une population : la taille, le poids, etc. ;
- le revenu des ménages d'une région ;
- la demande pour plusieurs produits ;
- certains indices boursiers;
- etc.

Paramètres de la loi normale

Afin de caractériser entièrement le modèle normal, nous devons préciser deux quantités, que l'on appelle ***paramètres***:

- la **moyenne** μ, indiquant la moyenne des observations, ou encore le centre de la cloche.

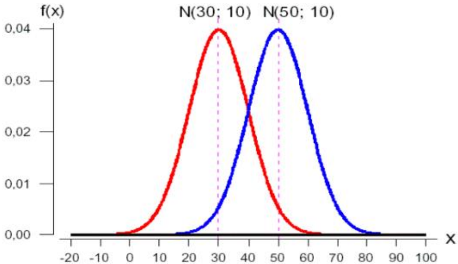
- l'**écart-type** σ, indiquant la dispersion des observations.

Notation:

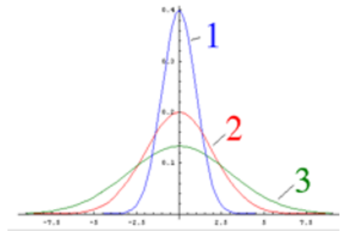
$$\boldsymbol{X \sim N(\mu, \sigma)}$$

Paramètres de la loi normale

Exemple de 2 lois normales avec le même écart-type



Trois courbes normales avec la même moyenne, mais respectivement un écart-type de 1, 2 et 3.



Cloche pointue	↔	Faible dispersion	↔	Petit écart-type
Cloche évasée	↔	Forte dispersion	↔	Grand écart-type

Propriétés de la loi normale

1) La densité de la normale est **unimodale** et **symétrique** par rapport à la moyenne.
Par conséquent, **mode = moyenne = médiane**.

2) **Une transformation linéaire** d'une variable normale est aussi **normale** :

Si $\boldsymbol{X \sim N(\mu, \sigma)}$ et $\boldsymbol{Y = aX + b}$ alors $\boldsymbol{Y \sim N(a\mu + b; |a|\sigma)}$

3) Si $\boldsymbol{X \sim N(\mu, \sigma)}$ et $\boldsymbol{Z = \frac{X-\mu}{\sigma}}$ alors $\boldsymbol{Z \sim N(0; 1)}$

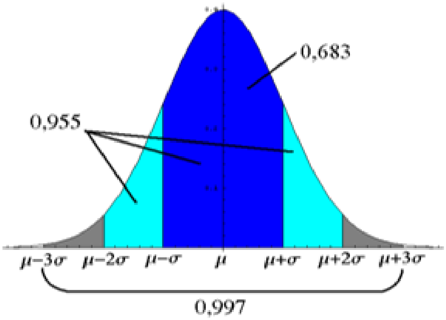
La loi **N(0, 1)** s'appelle la **loi normale centrée réduite** ou **la loi normale standard.**

Propriété empirique de la loi normale

Peu importe les valeurs des paramètres μ et σ, les proportions suivantes sont **toujours** respectées pour une variable de loi normale :

- environ **68,3%** des observations se situent à moins de **1** écart-type de la moyenne;
- environ **95,5%** des observations se situent à moins de **2** écarts-types de la moyenne;
- environ **99,7%** des observations se situent à moins de **3** écarts-types de la moyenne.

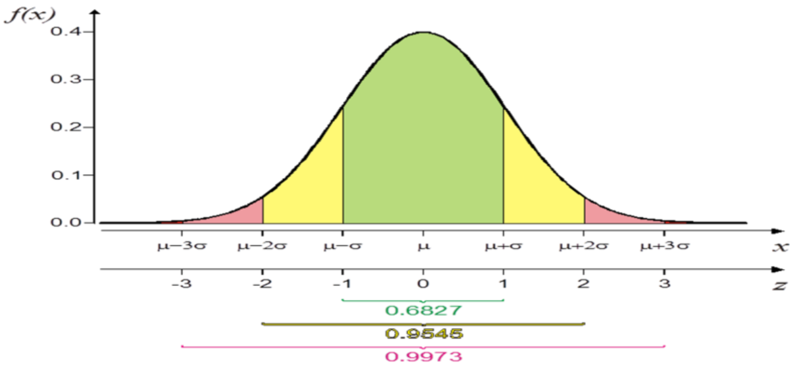
Par exemple, si la note à l'examen final suit une loi normale de moyenne 70% avec un écart-type de 10%. Alors, on s'attend à ce qu'environ 95.5% des étudiants aient une note entre 50% $(70 - 2 \times 10)$ *et 90%* $(70 + 2 \times 10)$*.*



Propriété empirique de la loi normale

Similairement, peu importe les valeurs des paramètres μ et σ, les proportions suivantes sont **toujours** respectées pour une variable de loi normale :

- environ **95%** des observations se situent à moins de **1,96** écart-type de la moyenne;
- environ **99%** des observations se situent à moins de **2,58** écarts-types de la moyenne;



Calcul de probabilités pour la normale

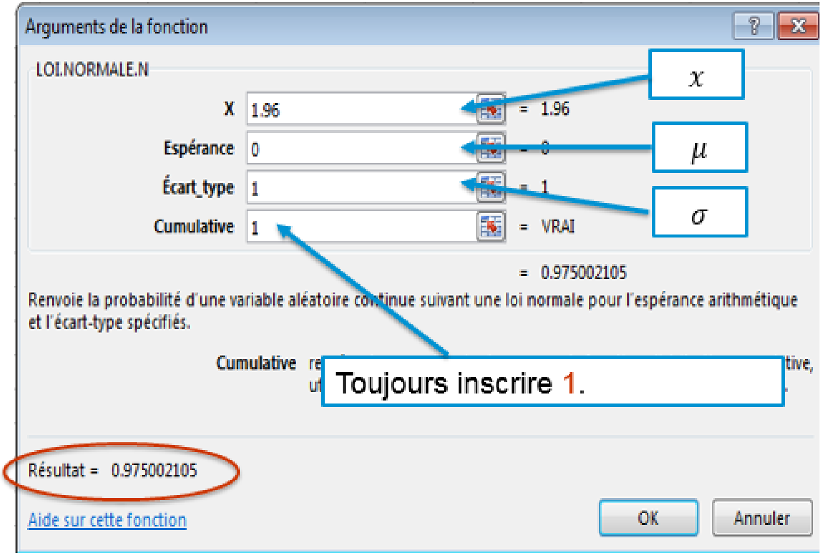
On peut calculer toutes les probabilités de la forme $\boldsymbol{P(X \leq x)}$ en utilisant Excel. Voir la diapositive qui suit. On peut en déduire toutes les autres probabilités utiles.

<u>Remarque importante</u>:
Contrairement à la loi binomiale, qui est de type discret, la loi normale est de type **<u>continu</u>** (une variable normale <u>peut prendre n'importe quelle valeur réelle</u>). Pour ce type de variable, on a $P(X = x) = 0$, quel que soit le nombre x considéré. Voici quelques conséquences:

- <u>Tous</u> nos calculs seront basés sur des probabilités du type $P(X \leq x)$, que nous obtiendrons avec Excel (jamais $P(X = x)$).
- On a toujours $P(X \leq x) = P(X < x)$, ce qui simplifie les calculs. En effet, on n'a pas à ce soucier du sens strict ou non strict des inégalités. Notons aussi que $P(X \geq x) = 1 - P(X < x)$.
- **C'est en fait l'aire sous la courbe normale, et non pas sa hauteur en chaque point, qui correspond à une probabilité.**

Utilisation d'Excel

À l'aide de la fonction Excel LOI.NORMALE.N, on peut calculer $P(X \leq x)$ lorsque $X \sim N(\mu, \sigma)$.



Utilité et importance de la loi normale

La loi normale est la distribution la **plus utilisée en pratique**, notamment parce que:

- elle décrit adéquatement les fluctuations de grand nombre de variables rencontrées en pratique.
- c'est un modèle flexible, permettant n'importe quelles valeurs pour la moyenne et l'écart-type.
- elle possède de belles propriétés, si bien qu'elle a donné lieu à une riche théorie mathématique, dont les retombées sont nombreuses et fort utiles dans la pratique.

Exemple

La durée de vie des transmissions automobiles AAA est une variable aléatoire distribuée selon une loi normale de moyenne 150 000 km et d'écart type 30 000 km.

a) Quelle est la probabilité qu'une transmission dure au moins 150 000 km?

b) Quelle est la probabilité qu'une transmission dure plus de 200 000 km ?

c) Quelle est la probabilité qu'une transmission dure entre 120 000 km et 180 000 km?

d) Le fabricant de la transmission ne veut absolument pas réparer sous garantie plus de 20% des transmissions. Quelle doit être, en kilomètres, la période de garantie?

Propriétés de la loi normale (suite)

4) La somme de variables aléatoires normales indépendantes est aussi normale

Soient $\boldsymbol{X_1}$, $\boldsymbol{X_2}$, …, $\boldsymbol{X_n}$, n variables aléatoires ***indépendantes*** telles que:

$$\boldsymbol{Xi} \sim N(\mu_i\,,\ \sigma_i\,)\ pour\ i\ =\ 1,\ \ldots,\ n$$

Si $\boldsymbol{Y}\ =\ \boldsymbol{X_1} +\ \boldsymbol{X_2}\ +\ \ldots\ +\ \boldsymbol{X_n}$, alors $\boldsymbol{Y} \sim \boldsymbol{N}(\boldsymbol{\mu}\,,\boldsymbol{\sigma})$ où

$$\boldsymbol{\mu} = \mu_1 + \mu_2 + \cdots + \mu_n\ \ et\ \ \boldsymbol{\sigma} = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}$$

Cas Gestion de projet

L'équipe en charge de la gestion d'un important projet doit faire des prévisions budgétaires. Le projet sera réalisé en deux phases successives (la 2e phase peut débuter seulement lorsque la 1ère est terminée). Chaque phase comporte un coût fixe et un coût variant selon sa durée :

Phase 1 : **100** milliers de \$ + **9** milliers de \$ par jour

Phase 2 : **75** milliers de \$ + **12** milliers de \$ par jour

Bien que l'on s'attende à ce que la première phase dure 50 jours et la deuxième 30 jours, il est difficile au stade actuel de faire des prévisions précises quant à la durée exacte des deux phases. La durée de chaque phase est donc incertaine (aléatoire), ce qui engendre de l'incertitude sur la durée du projet et son coût.

Durée aléatoire des phases

Supposons que selon l'expérience du passé pour ce type de projet, il soit plausible de croire que les variations de durées soient adéquatement décrites par des lois normales, dont les paramètres sont fournis dans le tableau suivant:

	moyenne	écart-type
D_1 (durée phase 1)	50	10
D_2 (durée phase 2)	30	4

Puisque les durées des phases sont incertaines (aléatoires) et que le coût du projet dépend de ces durées, il en résulte que **le coût total** du projet est également incertain (aléatoire)?

Comment bien saisir le comportement aléatoire du coût?

Comment bien quantifier le risque?

Modélisation

La première étape pour analyser cette situation est de faire l'inventaire des composantes (paramètres, mesure de performance, etc.). Dans le cas présent, les composantes sont :

- Durée de la phase 1, notée D_1
- Durée de la phase 2, notée D_2.
- Coût fixe de la phase 1, noté F_1 pour la suite
- Coût fixe de la phase 2, noté F_2 pour la suite
- Coût variable de la phase 1, noté V_1 pour la suite
- Coût variable de la phase 2, noté V_2 pour la suite
- Coût total de la phase 1, noté X_1pour la suite
- Coût total de la phase 2, noté X_2pour la suite
- Coût total du projet (**mesure de performance**), noté *Y* pour la suite.

On modélise ensuite la relation entre ces composantes.

Cas Gestion de projet

Question 1

Quelle est la loi de probabilité qui décrit le comportement aléatoire du <u>coût total</u> du projet? Quel est le coût moyen (ou espéré) du projet?

Indice : utilisez les propriétés 2 et 4 de la loi normale.

Question 2

Estimez la probabilité que le coût réel dépasse 1 million de dollars.

Question 3

Estimez la réserve monétaire qui doit être constituée pour être à 90% certain de couvrir les dépenses.