Modèles d'aide à la décision 4-600-04

Séance 2
Simulation de Monte-Carlo

Plan

- 2.1 Introduction
- 2.2 Rappel sur les probabilités et les statistiques
- 2.3 Implantation d'un modèle de simulation de Monte-Carlo
- 2.4 Présentation des résultats d'un modèle de simulation de Monte-Carlo
- 2.5 Exemple 2

2.1 Introduction

- Dans plusieurs modèles, il arrive que la valeur d'une ou de plusieurs variables indépendantes (ou paramètres) soit inconnue ou incertaine.
- Dans ce cas, il existe un niveau d'incertitude quant à la valeur de certaines variables dépendantes :

$$Y = f(X_1, X_2, \ldots, X_k)$$

- La simulation peut être utilisée pour modéliser ce genre de situations.

2.1 Variables aléatoires et risque

- Une *variable aléatoire* est une variable dont la valeur ne peut être prédite ou fixée avec certitude.
- Plusieurs cellules d'entrée dans les modèles de tableurs sont en fait des variables aléatoires.
  - Le coût futur de la matière première
  - Les taux d'intérêt futurs
  - Le nombre futur d'employés dans une entreprise
  - La demande prévue d'un produit
- Les décisions prises dans un contexte d'incertitude impliquent souvent un risque.
- Le risque implique un potentiel de perte.

2.1 Pourquoi analyser le risque?

- Une possibilité consiste à remplacer la variable par son espérance.
- Cela ne nous donne aucune information sur la variabilité de la fonction de mesure de performance sur laquelle nous basons notre décision.
- Supposons qu'un investissement de 1 000$ procure un gain espéré de 10 000$ dans deux ans. Allez-vous investir si...
  - Le gain peut varier entre 9 000$ et 11 000$?
  - Le gain peut varier entre -30 000$ et 50 000$?
- Des alternatives avec la même valeur espérée peuvent engendrer des niveaux de risque différents.

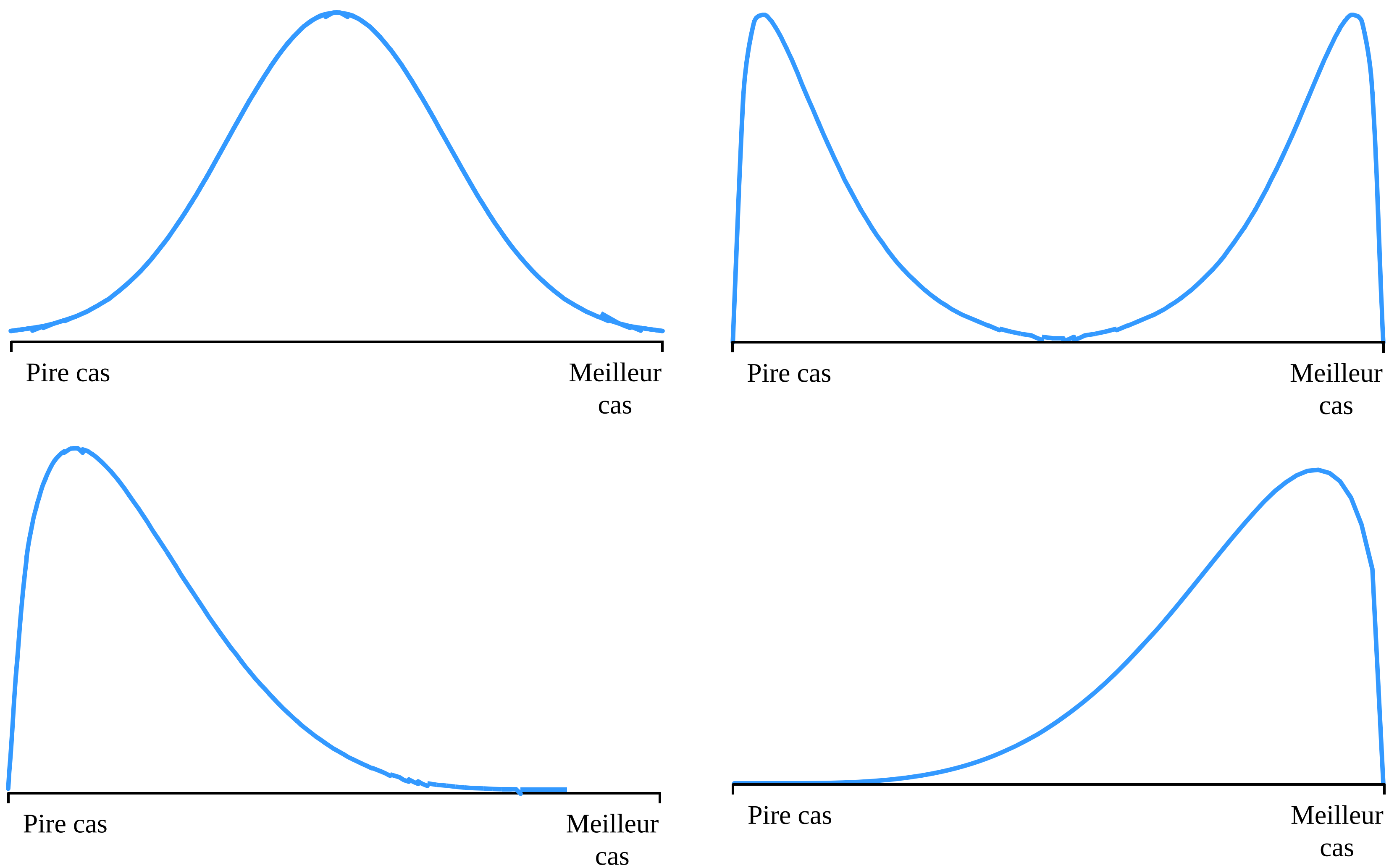
2.1 Méthodes d'analyse de risque

- Fonctions d'utilité (voir section 14.15 de Ragsdale)
- Analyse du meilleur cas / pire cas.
- Analyse de scénarios
- Simulation de Monte-Carlo

2.1 Analyse du meilleur cas / pire cas

- Meilleur cas : mettre les valeurs les plus optimistes pour tous les paramètres incertains.
- Pire cas : mettre les valeurs les plus pessimistes pour tous les paramètres incertains.
- Très facile à faire mais ne fournit aucune information sur la *distribution* des résultats possibles à l'intérieur de ces deux extrêmes.

2.1 Analyse du meilleur cas / pire cas



2.1 Analyse de scénarios

- Mettre des valeurs différentes pour les cellules (paramètres) incertaines et vérifier le résultat.
- Facile à faire avec un “bon” modèle fait sur un tableur.
- Défauts :
  - Les valeurs peuvent être choisies de façon biaisée.
  - Des centaines ou des milliers de scénarios peuvent être requis pour générer une distribution représentative.

2.1 Simulation de Monte-Carlo

- Analyse de scénarios automatisée.
- Valeurs des cellules sont choisies d'une façon non biaisée.
- L'ordinateur génère des centaines (voire des milliers) de scénarios.
- Nous analysons les résultats de ces scénarios de manière à mieux comprendre le comportement de la mesure de performance.
- Ça nous permet de prendre des décisions basées sur une solide évidence empirique.

2.1 Exemple 1 : Restaurant chez Joe

- Depuis 2 semaines Joe offre une nouvelle pâtisserie fraiche du jour sur son menu.
- Chaque matin, il prépare 15 de ces pâtisseries.
- Chaque soir il jette les invendues.
- Le coût de préparation est de 6$/p.
- Le prix de vente est de 10$/p.
- On a des statistiques pour les 2 semaines. (voir fichier 4600-2.1-Restaurant_Joe-Data.xlsx)

2.1 Exemple 1 : Restaurant chez Joe

- Le nombre de clients est uniformément réparti entre 1 et 6.
- Le nombre de portions/clients est 1 ou 2
- La taille de l'échantillon est assez petite:
  - On n'a pas tous les scénarios possibles;
  - Tous les scénarios n'ont pas la même probabilité de se réaliser.
- On pourrait faire des profits si Joe préparait moins de pâtisseries chaque matin… Mais combien?

2.1 Exemple 1 : Restaurant chez Joe

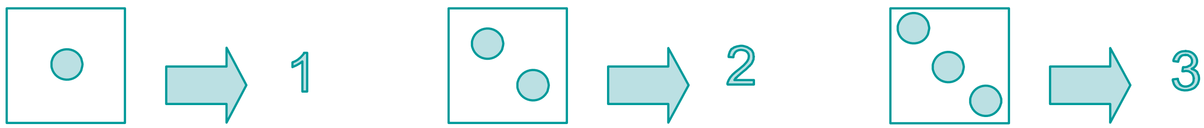
- Dans ce modèle on a des:
  - Paramètres certains :
    - Coût de production (6$/p.)
    - Prix de vente (10$/p.)
  - Variable de décision:
    - Nombre de pâtisseries préparées chaque jour
  - Paramètres incertains (variables aléatoires) :
    - Nombre de clients
    - Nombre de portions par clients
- Dans une simulation, on peut générer des réalisations de ces variables aléatoires pour obtenir la réalisation d'un scénario…. On peut en générer 100, 1 000, 10 000, 100 000, etc. et récolter des statistiques réalistes et utiles.

2.2 Rappel sur les probabilités et sur les variables aléatoires

- Une variable aléatoire est une variable qui prendra une valeur numérique, mais on ne peut savoir à l'avance quelle sera cette valeur.

- On distingue 2 grandes classes de variables aléatoires : les discrètes et les continues.
  - Les variables discrètes: les valeurs possibles sont clairement distinctes les unes des autres.
  - Les variables continues: les valeurs possibles peuvent être très rapprochées…

2.2 Variables aléatoires discrètes

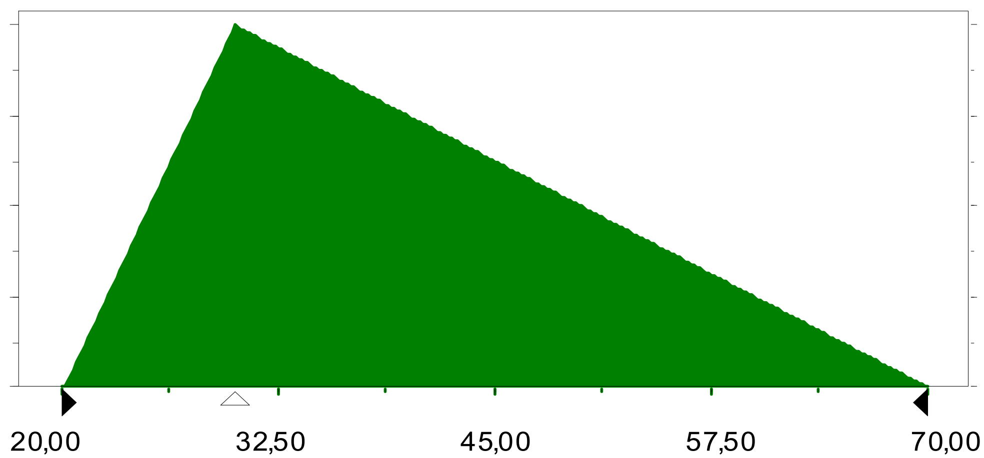


- X : nombre de points sur le dé
- Distribution de probabilité :

x	1	2	3	4	5	6
$P(X=x)$	1/6	1/6	1/6	1/6	1/6	1/6

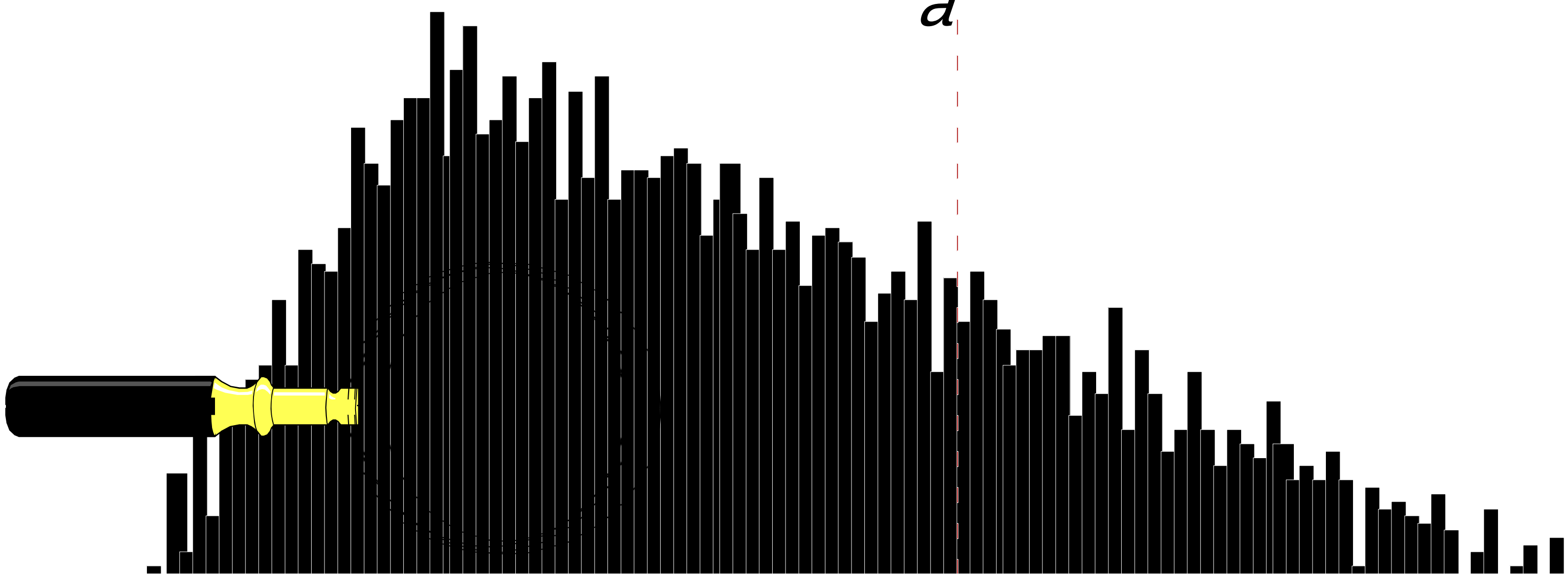
2.2 Variables aléatoires continues

- X : Demande (valeur quelconque entre 20 et 70, le plus souvent autour de 30)
- Densité de probabilité :



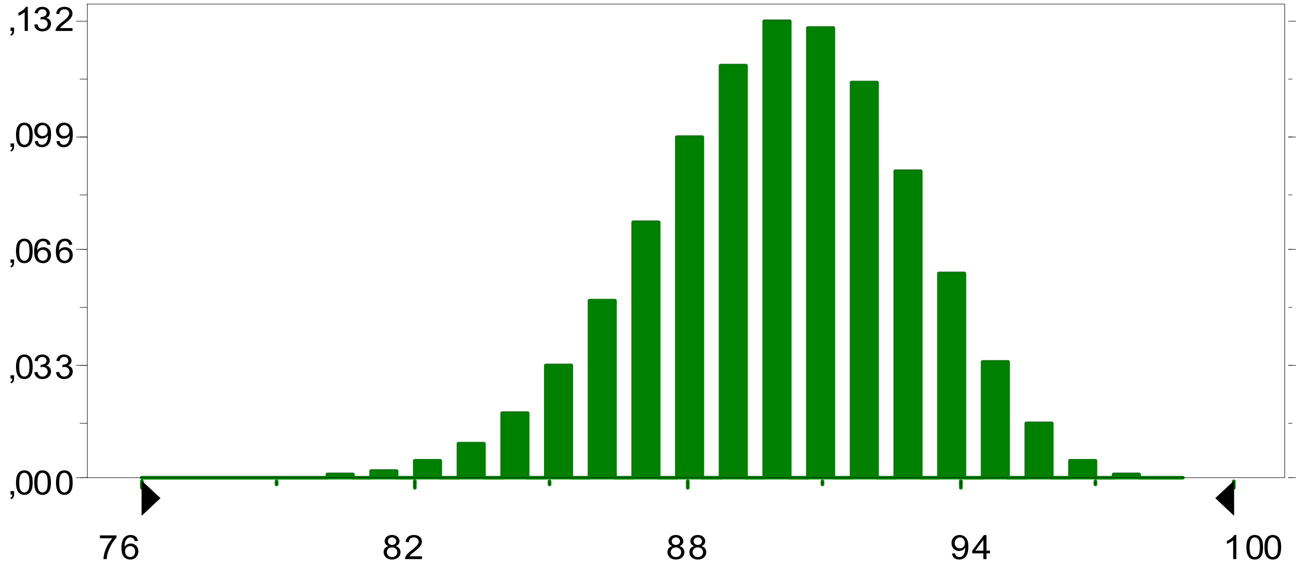
2.2 Distribution de probabilité

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



2.2 Distribution Binomiale

- Distribution discrète utilisée pour modéliser un nombre de succès parmi un nombre d'expériences.
- Paramètres : n = nb d'exp. et p= proba. de succès.
- Ex : Nombre de clients satisfaits parmi 100, nombre de produits défectueux parmi 1000, etc.
- Illustration de la fonction de probabilité où n=100 et p=0,9

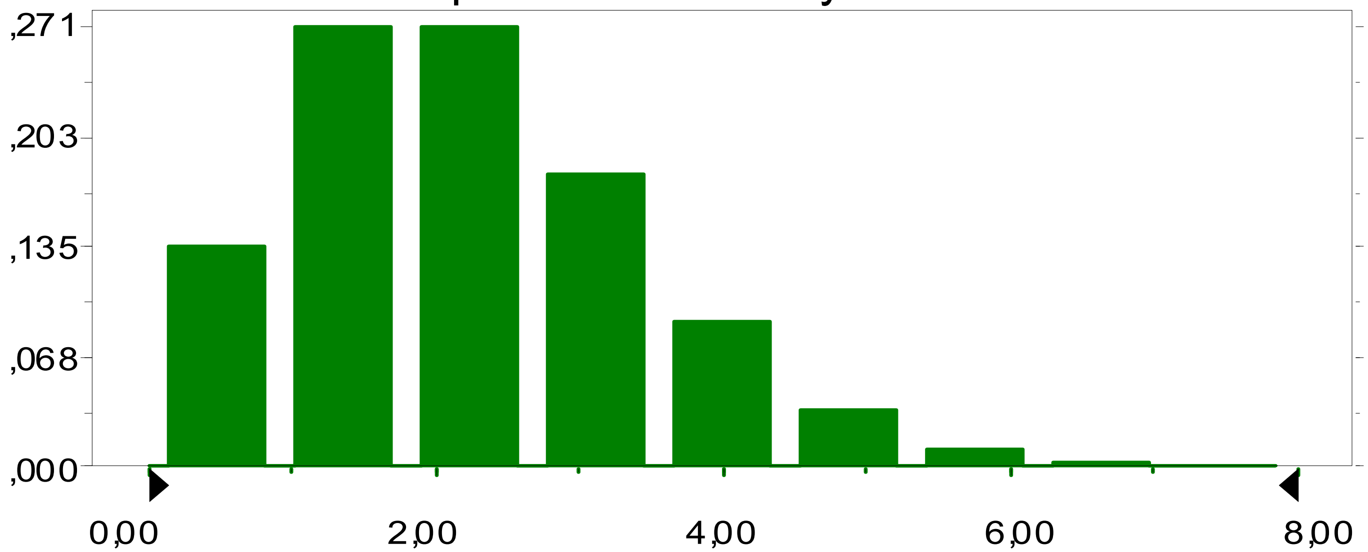


2.2 Distribution Binomiale

- Valeur la plus fréquente : $n{\cdot}p$
- C'est aussi la valeur espérée.
- Les valeurs sont toutes entre 0 et n.
- Les valeurs inférieures ou supérieures (*de loin*) à $n{\cdot}p$ sont très peu probables.

2.2 Distribution de Poisson

- Distribution discrète utilisée pour modéliser un nombre de réalisations par unité de mesure.
- Paramètres nécessaires : moyenne
- Ex : Nombre de pannes durant un mois, nombre d'items demandés par client, nombre de produits défectueux par lot, etc.
- Illustration de la fonction de probabilité où moyenne=2

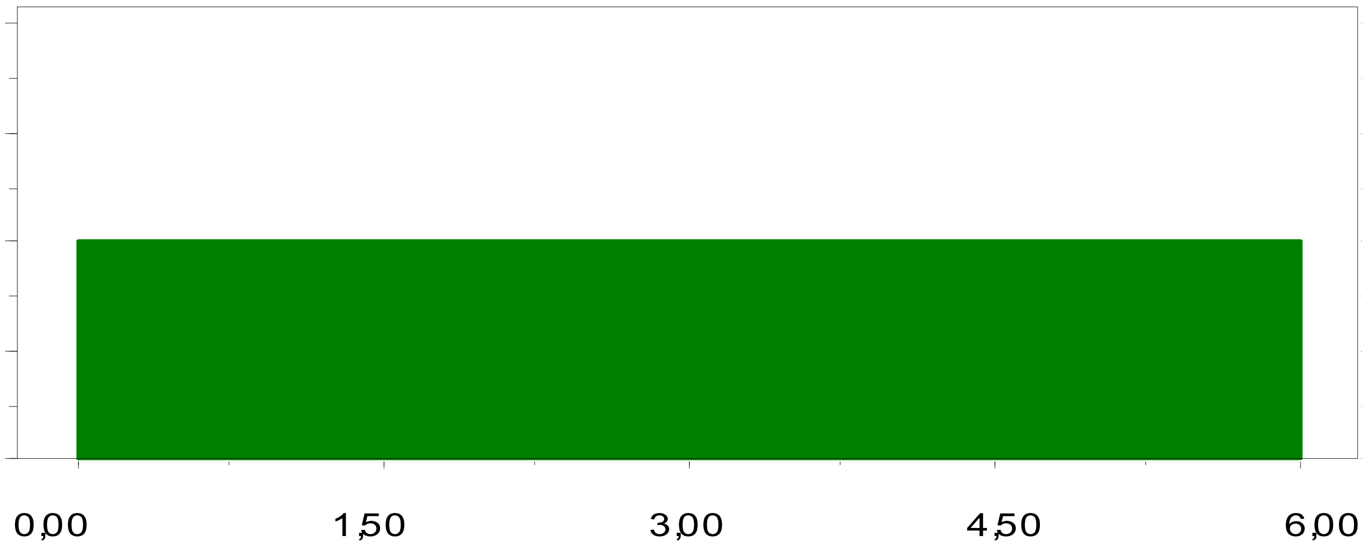


2.2 Distribution uniforme discrète

- Distribution modélisant une variable aléatoire pour laquelle toutes les valeurs, entre un minimum et un maximum, sont équiprobables.
- Paramètres nécessaires: valeurs minimale et maximale.
- Ex : Un dé (Voir plus tôt), la loterie, les cartes, le nombre de clients chez Joe.

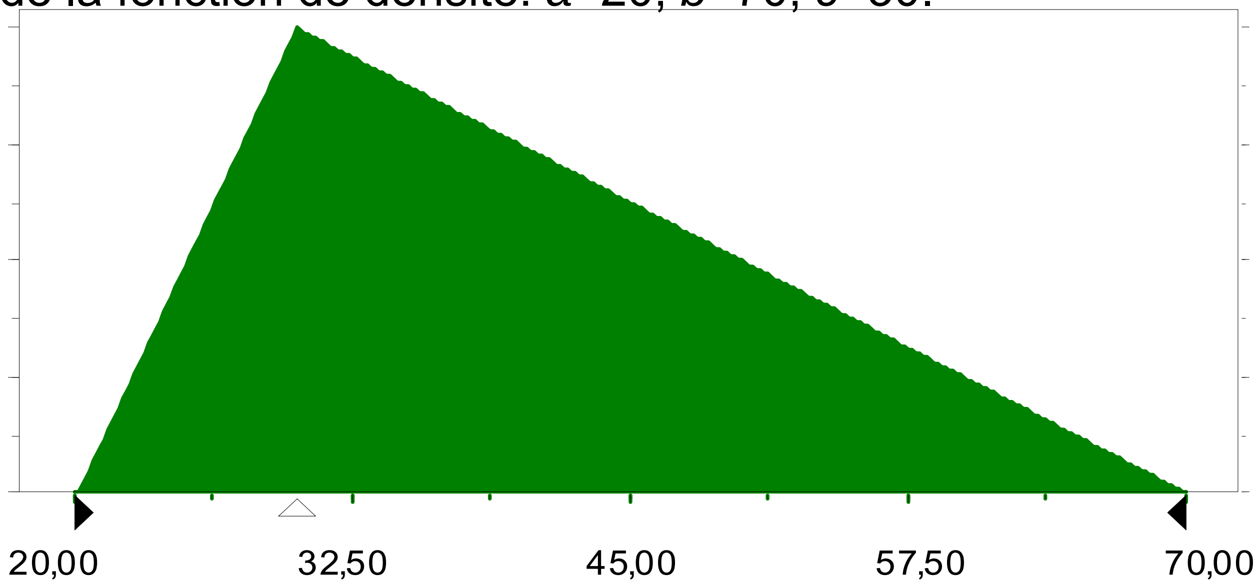
2.2 Distribution uniforme continue

- Distribution continue modélisant une variable aléatoire pour laquelle chaque valeur, entre un minimum et un maximum, est équiprobable.
- Paramètres : *a*=minimum et *b*=maximum.
- Ex : Délai d'attente, etc.
- Illustration de la fonction de densité : *a*=0 et *b*=6.



2.2 Distribution triangulaire

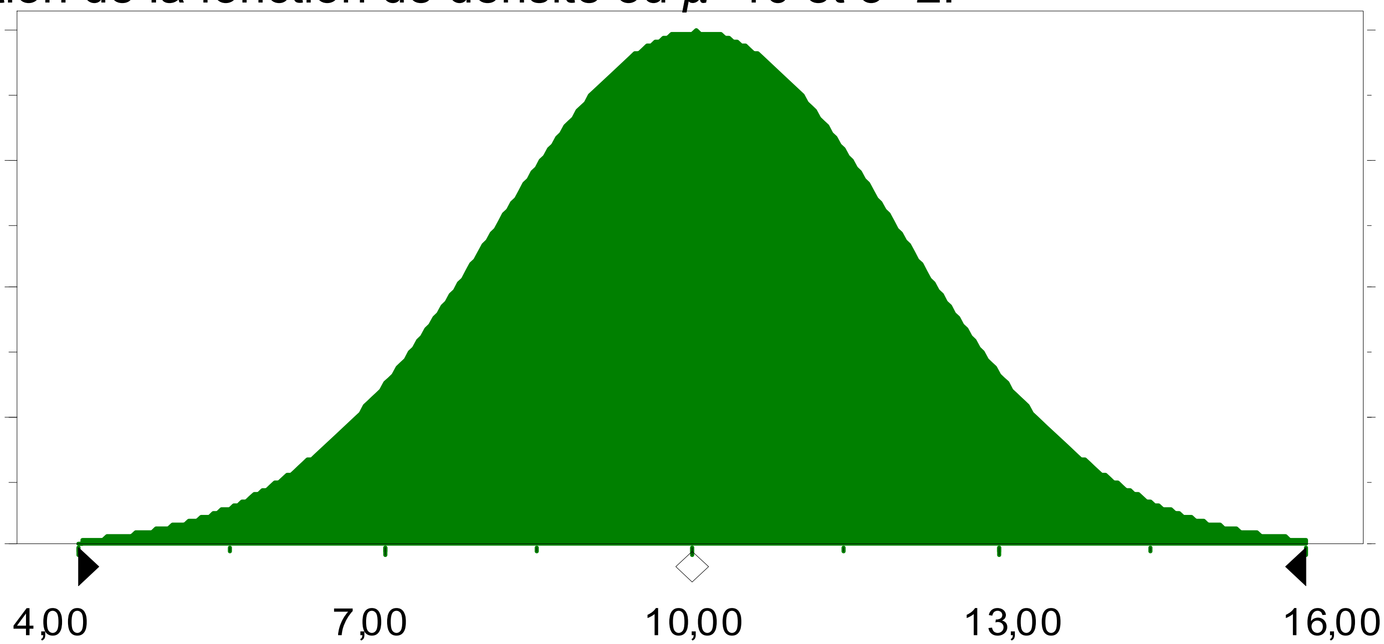
- Distribution continue modélisant une variable aléatoire dont on connaît que le minimum, le maximum et la valeur la plus probable.
- Paramètres : *a*=min., *b*=max. et *c*=la plus probable.
- Ex : Durée d'une tâche, coût unitaire de fabrication, etc.
- Illustration de la fonction de densité: *a*=20, *b*=70, *c*=30.



HEC MONTRÉAL

2.2 Distribution normale

- Distribution **symétrique** continue (forme de cloche).
- Paramètres : μ=moyenne et σ=écart-type.
- Ex : Coûts d'opération, temps d'opération de certains procédés, taux de rendement d'une action, etc.
- Illustration de la fonction de densité où μ=10 et σ=2.



2.2 Distribution normale

- Moyenne :

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$

- Variance : moyenne des carrés des écarts à la moyenne

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N} (x_i - \mu)^2$$

- Écart-type : Écart-type : racine carré de la variance (même unité de mesure que la moyenne).

2.3 Simulation de Monte-Carlo

Les étapes d’une simulation de Monte-Carlo :

1. Obtenir les distributions de probabilités des paramètres incertains (information statistique ou analyse).
2. Construire un modèle flexible permettant de calculer la valeur des indicateurs économiques pour une réplication de valeurs de ces paramètres.
3. Générer une série de réplications des paramètres incertains en générant des nombres aléatoires selon les distributions de probabilités trouvées. Calculer les indicateurs économiques pour chacune de ces réplications.
4. Présenter les résultats sous la forme d’une distribution de probabilité.

2.3 Simulation de Monte-Carlo

- Une «réplication» correspond à la génération de valeurs aléatoires pour tous les paramètres incertains.
- Pour chaque réplication spécifique, un «résultat» est calculé, à partir du modèle déterministe. Ce résultat est noté afin de produire une distribution de tous les résultats des réplications.
- Ce résultat sera différent pour chaque réplication. Une étude statistique permet d'étudier le comportement du « résultat ».
- Une étude statistique comprend des tableaux, graphiques, statistiques descriptives et intervalles de confiance (voir séance 4).

2.3 Générateurs de nombres aléatoires

- Un générateur de nombres aléatoires est une fonction mathématique qui génère aléatoirement une valeur à partir d'une distribution particulière de probabilité.
- Nous pouvons implanter des générateurs de nombres aléatoires pour les cellules incertaines à l'aide de Risk Solver Platform.

2.3 Risk Solver Platform

- Risk Solver Platform est une macro complémentaire (add-in) qui simplifie l'implantation d'un modèle de simulation sur un tableur.
- RSP fournit :
  - fonctions pour générer des nombres aléatoires
  - commandes pour exécuter la simulation
  - rapports graphiques et statistiques des données de la simulation

2.3 Risk Solver Platform

- Aide sur l'utilisation de RSP :
  - Sections 12.5 à 12.9 de Ragsdale (6^e édition)
  - Tutoriels et Vidéos disponibles en ligne (voir les liens dans le menu « Help » de RSP)
  - Guides (voir le menu « Help » de RSP)
  - Séance 2 d'initiation à Excel
  - Séance de TP 1

2.3 Exemple 1 : restaurant Joe

- Rappel : paramètres incertains (variables aléatoires) :
  - Nombre de clients
  - Nombre de portions par clients
- Quel type de distribution de probabilités peut être utilisé pour modéliser chacun de ces paramètres?
- On commencera par établir un modèle permettant d'analyser le profit quotidien de Joe en supposant que Joe décide de faire 15 patisseries à chaque jour. Le fichier 4600-2.3-Restaurant_Joe-Sol.xlsx présente le modèle et les résultats d'une simulation à 1000 réplications à l'aide de RSP.

2.3 Résumé des principales fonctions de RSP

- **Cellule** (variable) **aléatoire** : se mettre dans la cellule et cliquer sur le Menu « Distributions » de RSP ou utiliser les fonctions Excel débutant par Psi (ex : = PsiNormal())
- **Cellule Output** (indicateur économique) : se mettre dans la cellule et cliquer successivement sur Results->Output->In Cell
- Fixer au besoin le nombre de réplications (Trials per simulation) et la graine du générateur (Sim. Random Seed) dans les options
- **Lancer la simulation** : Menu Simulate-> Run Once

HEC MONTRÉAL

2.3 Résumé des principales fonctions de RSP

- **Rapport sommaire** : après la simulation, appuyer successivement sur Reports->Simulation.
- **Résultats sur une cellule output** : après la simulation, double-cliquer sur la cellule.
- **Statistiques sur une cellule :** avant la simulation, utiliser une des fonctions de RSP (ex : psimean(), psimax(),…). Voir par exemple les cellules B11 à B14 de la feuille « Modèle – RSP » du fichier 600-2.3-Restaurant_Joe-Sol.xlsx.

2.3 Résumé des principales fonctions de RSP

- **Histogramme** : utiliser la fonction PsiFrequency() et les graphiques en colonnes d'Excel. Voir par exemple la feuille « Graphique » du fichier 4600-2.3-Restaurant_Joe-Sol.xlsx.
- **Percentiles :** utiliser la fonction PsiPercentile(). Voir par exemple la feuille « Percentiles » du fichier 4600-2.3-Restaurant_Joe-Sol.xlsx.
- **Données complètes** : utiliser la fonction PsiData(). Voir par exemple la feuille « Échantillon généré » du fichier 4600-2.3-Restaurant_Joe-Sol.xlsx.

2.4 Présentation des résultats d'une simulation de Monte-Carlo

- Échantillon complet
- Fréquences (tableaux et histogrammes)
- Statistiques descriptives
- Centiles
- Intervalles de confiance (voir séance 4)
- Etc…

2.4 Statistiques descriptives

- Mesures de tendance centrale :
  - Moyenne : $$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i$$
  - Médiane : valeur centrale lorsque les données sont ordonnées en ordre croissant.
  - Mode : observation la plus fréquente.

2.4 Statistiques descriptives

- Mesures de dispersion :
  - Étendue : différence entre la valeur la plus grande et la valeur la plus petite.
  - Variance : moyenne des carrés des écarts à la moyenne.

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2$$

2.4 Statistiques descriptives

- Mesures de dispersion :
  - Écart-type : racine carré de la variance (même unité de mesure que la moyenne).
- Coefficient de variation : procure une mesure relative de la dispersion relative à la moyenne. Très utile pour comparer la variabilité de deux ensembles de données ayant une unité de mesure différente.

$$CV = \frac{\sigma}{\mu}$$

2.5 Exemple 2 : analyse du temps d'un procédé

- La compagnie MAD vient de recevoir un appel d'offre pour 6 pièces articulées servant dans le milieu minéral.
- La fabrication de ces pièces se fait en deux étapes successives sur deux machines différentes.
- Les dirigeants de MAD aimeraient évaluer le temps total qu'il faudra pour remplir cette commande et ainsi évaluer le coût relatif au personnel.

2.5 Exemple 2 : analyse du temps d'un procédé

- Le temps moyen nécessaire pour compléter la première étape est de 20 heures avec un écart-type de 4 heures.
- Pour la deuxième étape le temps moyen est de 21 heures avec un écart-type de 3 heures.
- Le premier réflexe a donc été de dire qu'en moyenne, il faudra 20 heures pour compléter la première étape sur la première pièce et 6 fois 21 heures pour compléter la deuxième étape sur toutes les pièces.

2.5 Exemple 2 : analyse du temps d'un procédé

- Il a donc été proposé de produire une soumission avec 146 heures de main-d'œuvre.
- Cette soumission comporte deux défauts :
  - Il s'agit d'une moyenne. Il est donc très probable que cela prenne plus de temps.
  - On ne tient pas compte des temps d'attente où la deuxième machine serait libre mais que la première étape n'est pas complétée. Cette situation est très probable!
- On désire concevoir un modèle pour décrire avec plus de précision le temps de ce procédé.

2.5 Exemple 2 : analyse du temps d'un procédé

- Quels sont les paramètres incertains (variables aléatoires)?
- Quel type de distribution de probabilités peut être utilisé pour modéliser chacun de ces paramètres?
- Le fichier 4600-2.5-Exemple2-Sol.xlsxprésente le modèle et les résultats d'une simulation à 1000 réplications à l'aide de RSP.