$$F_0(x) = \frac{\#(Y_i \le x)}{n}$$

2- Pour cet échantillon de données, supposons que la CDF théorique $F_e(x)$ (Expected CDF) pour une distribution uniforme soit définie comme $F_e(x)=x$ pour chaque valeur x dans l'intervalle [0,1].

- Déterminer $F_e(x)$
- Calculer la statistique K du test KS définie comme :

$$K = \max_x |F_e(x) - F_0(x)|$$

3- Comparer la statistique K avec la valeur critique correspondante pour un niveau de signification $\alpha=0.05$ (En se basant sur le Tableau KS ci-dessous).

$n \backslash \alpha$	0.001	0.01	0.02	0.05	0.1	0.15	0.2
1		0.99500	0.99000	0.97500	0.95000	0.92500	0.90000
2	0.97764	0.92930	0.90000	0.84189	0.77639	0.72614	0.68377
3	0.92063	0.82900	0.78456	0.70760	0.63604	0.59582	0.56481
4	0.85046	0.73421	0.68887	0.62394	0.56522	0.52476	0.49265
5	0.78137	0.66855	0.62718	0.56327	0.50945	0.47439	0.44697
6	0.72479	0.61660	0.57741	0.51926	0.46799	0.43526	0.41035
7	0.67930	0.57580	0.53844	0.48343	0.43607	0.40497	0.38145
8	0.64098	0.54180	0.50654	0.45427	0.40962	0.38062	0.35828
9	0.60846	0.51330	0.47960	0.43001	0.38746	0.36006	0.33907
10	0.58042	0.48895	0.45662	0.40925	0.36866	0.34250	0.32257
11	0.55588	0.46770	0.43670	0.39122	0.35242	0.32734	0.30826
12	0.53422	0.44905	0.41918	0.37543	0.33815	0.31408	0.29573
13	0.51490	0.43246	0.40362	0.36143	0.32548	0.30233	0.28466
14	0.49753	0.41760	0.38970	0.34890	0.31417	0.29181	0.27477
15	0.48182	0.40420	0.37713	0.33760	0.30397	0.28233	0.26585
16	0.46750	0.39200	0.36571	0.32733	0.29471	0.27372	0.25774
17	0.45440	0.38085	0.35528	0.31796	0.28627	0.26587	0.25035
18	0.44234	0.37063	0.34569	0.30936	0.27851	0.25867	0.24356
19	0.43119	0.36116	0.33685	0.30142	0.27135	0.25202	0.23731
20	0.42085	0.35240	0.32866	0.29407	0.26473	0.24587	0.23152

4- Quelle conclusion pouvez-vous tirer de cette comparaison ?

Exercice 2 : Inverse Transform Method - Geometric Variate

Dans le cadre d'une entreprise de commerce en ligne, on cherche à modéliser le nombre de tentatives nécessaires pour réaliser la première vente. Cette situation suit une distribution géométrique, avec une probabilité de succès de p=0.2 (soit 20%) pour chaque tentative.

Institut Supérieur d'Informatique et de Mathématiques de Monastir

ISIMM

Filière : L2 INFO	**Matière : System Modeling & Simulation**		**Enseignante : Nada Haj Messaoud**
Date : 18 / 01 / 2025	Nbr de Crédits : **3**	Coefficient : **1.5**	**Documents autorisés : Non**
Durée d'examen : 1h30	Régime d'évaluation : **Mixte** EX (70%) + DS (10%) + TP (20%)		**Nombre de pages : 07**

Examen – S1 – 2024/2025

- *Calculatrice autorisée.*
- *Lire attentivement les questions.*

Partie I : QCM (5 points)

1. Qu'est-ce que la modélisation et la simulation ?

A. Une méthode pour résoudre des équations complexes uniquement.
B. Une approche pour reproduire le comportement d'un système réel à l'aide de modèles mathématiques.
C. Une technique de cryptographie avancée.
D. Une méthode pour concevoir des systèmes matériels uniquement.

2. Laquelle des approches suivantes est considérée comme une approche analytique ?

A. Simulation.
B. Résolution d'équations mathématiques.
C. Observation d'un système réel en fonctionnement.
D. Génération de nombres aléatoires.

3. Quel est la première étape du processus de simulation ?

A. Analyser les résultats.
B. Construire un modèle mathématique.
C. Collecter des données statistiques.
D. Définir le problème et les objectifs.

4. Qu'est-ce qui n'est pas une composante clé d'un modèle de simulation ?

A. Les entrées.

B. Les sorties.

C. La base de données utilisateur.

D. Les relations entre variables.

5. La méthode Monte Carlo repose sur :

A. L'échantillonnage aléatoire pour résoudre des problèmes numériques.

B. Des calculs déterministes.

C. L'observation directe d'un système réel.

D. La construction de modèles physiques pour tester des hypothèses.

6. Le test Chi-Square est principalement utilisé pour :

A. Vérifier si une séquence suit une distribution théorique donnée.

B. Comparer deux distributions continues.

C. Générer des nombres pseudo-aléatoires.

D. Détecter les dépendances entre les nombres générés.

7. Le test Kolmogorov-Smirnov (KS) compare :

A. L'indépendance des nombres générés.

B. La fréquence de répétition des nombres générés.

C. La relation linéaire entre deux variables.

D. La CDF observée (F0(x)) à la CDF attendue (Fe(x)) pour évaluer l'adéquation des données à la distribution théorique.

8. Quel est le rôle principal d'un générateur de congruence linéaire (LCG) ?

A. Générer des nombres parfaitement aléatoires.

B. Tester les hypothèses statistiques.

C. Générer des séquences pseudo-aléatoires.

D. Calculer les valeurs de fonction gamma.

9. Dans la formule d'un LCG, que représente m ?

A. La graine initiale.

B. Le multiplicateur.

C. Le module, définissant la plage des valeurs générées.

D. Le décalage.

10. Quelle est une limitation principale d'un générateur de congruence linéaire (LCG) ?

A. Il est très lent à exécuter.

B. Il génère des séquences déterministes, donc pas parfaitement aléatoires.

C. Il ne peut pas fonctionner avec de grands modules mmm.

D. Il nécessite des tests coûteux pour être initialisé.

Partie II (3 points)

1- Quelle est la différence entre « Single server infinite population » et « Single server finite population ».

2- Quelle est la différence principale entre les approches « Time-Stepping » et « Event-Scheduling ».

3- Une usine de fabrication de téléphones mobiles utilise un système de production en **trois étapes** successives pour assembler chaque téléphone. Chaque téléphone passe par les trois étapes suivantes :

- **Étape 1 : Assemblage des composants**
- **Étape 2 : Installation des logiciels**
- **Étape 3 : Test de qualité**

- Comment peut-on modéliser ce système ? Expliquez pourquoi ?

Partie III (12 points)

<u>Exercice 1 : Test Kolmogorov-Smirnov (KS) pour une distribution uniforme</u>

Imaginons que nous avons un échantillon de 10 valeurs qui sont les suivantes (tirées aléatoirement) :

$$X= [0.12, 0.35, 0.55, 0.72, 0.88, 0.42, 0.69, 0.21, 0.80, 0.61]$$

Nous voulons tester si cet échantillon de données suit une distribution uniforme sur l'intervalle [0, 1].

1- Calculer la CDF observée $F_0(x)$ (Observed CDF) en utilisant la formule :

A partir des six valeurs aléatoires uniformes U générées :

$U_1=0.35, U_2=0.52, U_3=0.08, U_4=0.90, U_5=0.65, U_6=0.75$

1- Calculer les valeurs de X pour chaque valeur de U en utilisant la formule inverse définie comme suit :

$$X = \left\lceil \frac{\ln(u)}{\ln(1-p)} \right\rceil$$

2- Calculer la moyenne du nombre d'essais nécessaires pour obtenir la première vente.

3- Déterminer la moyenne théorique pour cette distribution géométrique qui est donnée par la formule 1/p.

4- Comparer la moyenne trouvée avec la moyenne théorique et interpréter les résultats.

Exercice 3 : Temps de trajet en taxi - Empirical Distribution

Vous êtes analyste pour une société de taxis, et vous souhaitez analyser les temps de trajet des clients dans une ville. Voici les données collectées pour les trajets de 150 clients, regroupés en intervalles de minutes :

Intervalle (Minutes)	Fréquence	Fréquence relative	Fréquence cumulée	Pente
$0 < x \leq 3$	10	?	?	?
$3 < x \leq 6$	30	?	?	?
$6 < x \leq 9$	40	?	?	?
$9 < x \leq 12$	50	?	?	?
$12 < x \leq 15$	20	?	?	?

1- Compléter le tableau :

- Calculer la fréquence relative pour chaque intervalle qui est la division de fréquence par le total des fréquences.
- Calculer la fréquence cumulée pour chaque intervalle : La fréquence cumulée représente le pourcentage des trajets dont le temps est inférieur ou égal à la borne supérieure de l'intervalle.

$$\text{Fréquence cumulée}_i = \sum_{j=1}^{i} \text{Fréquence}_j$$

- Calculer « Slope » pour chaque intervalle :

$$\alpha_i = \frac{c_i - c_{i-1}}{a_i - a_{i-1}}$$

2- Tracer l'histogramme de la distribution empirique.

3- Calculer le temps de trajet x correspondant à une fréquence cumulée de 0.8, càd 80% des trajets ont une durée inférieure ou égale à x.

$$x = a_{i-1} + \frac{u - c_{i-1}}{\alpha_i}$$

Exercice 4 : Simulation d'un Parking avec un maximum de 3 places

Vous êtes responsable de la gestion d'un petit parking situé à proximité d'un bureau de poste (ou d'une banque). Ce parking dispose d'un maximum de **3 places de stationnement**. Le nombre de places est limité, et les clients (voitures) arrivent à des moments aléatoires pour trouver une place. Une fois qu'un client trouve une place, il y stationne pendant un certain temps avant de repartir.

Temps inter-arrivées	0	5	2	7	4	6	3
Temps de stationnement	6	8	10	4	9	5	7

Composants du système :

- **État du système :**
  - **LQ(t)** : Nombre de voitures en attente à t (si le parking est plein).
  - **LS(t)** : Nombre de places occupées à t ($0 \leq LS(t) \leq 3$).
- **Entités :**
  - **Les voitures** sont les clients qui arrivent au parking et stationnent pendant un certain temps.
  - **Le parking** est un espace limité avec seulement 3 places disponibles.
- **Événements :**
  - **A(t)** : Arrivée d'un client au temps t. Si une place est disponible, le client occupe une place. Sinon, il attend.
  - **D(t)** : Départ d'un client au temps t. Une voiture quitte le parking, libérant une place pour un autre client.
  - **E (480)**: Fin de la simulation au temps t=480 (8 heures de simulation).

- **Notifications d'événements :**
  - **(A,t,C_i)**: Le client C_i arrive au temps t.
  - **(D,t,C_i)** : Le client C_i quitte le parking au temps t.
  - **(E,480)** : La simulation s'arrête au temps t=480.

Statistiques à collecter :

1. **S** : Somme des temps de réponse des clients ayant quitté le parking jusqu'à l'heure actuelle.
2. **F** : Nombre de voitures ayant passé **5 minutes ou plus** dans le parking.
3. **N_D** : Nombre total de départs avant la fin de la simulation.

Travail demandé : Dessiner le tableau de simulation pour obtenir les résultats demandés.