

TP 2 : Variables discrètes

A. PHILIPPE

Exercice 1 Autour de la loi binomiale

Soit X une variable aléatoire. On suppose que la loi de X est la loi binomiale de paramètres (n,p) où $n \in \mathbb{N}$ et $p \in [0,1]$.

On veut calculer

→ les probabilités $P(X=k)$ pour tout $k \in \{0, \dots, n\}$.

On peut ensuite représenter la loi de X (i.e. $P(X=k)$ en fonction de k) par un diagramme bâton

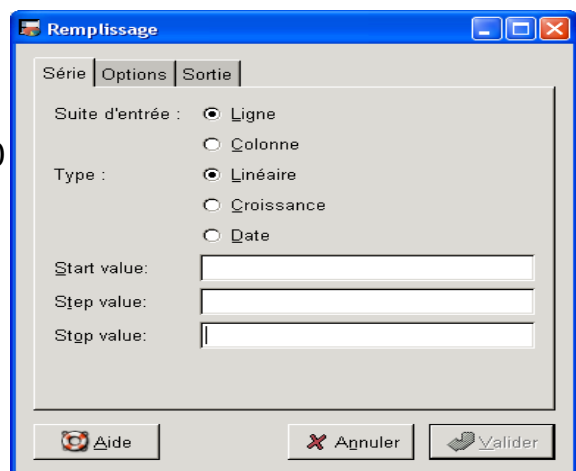


→ la fonction de répartition, c'est à dire $F(k) = P(X \leq k)$

(1) Calculer les probabilités $P(X=k)$ lorsque la loi de X est la loi binomiale de paramètres $n=15$ et $p=0,3$.

Indications :

- la colonne A contient les entiers de 0 à 15. Pour remplir la colonne:
 - sélectionner la case A1
 - Aller dans le menu **édition/remplir/séries**
 - Compléter la fenêtre **Remplissage**



- La colonne B contient les probabilités $P(X=k)$. On les calcule à l'aide de la fonction **Binomdist** (voir **Illustration 1**)
- Le paramètre cumulatif prend la valeur FAUX (pour les probabilités $P(X=k)$) ou VRAI (pour la probabilités $P(X \leq k)$)

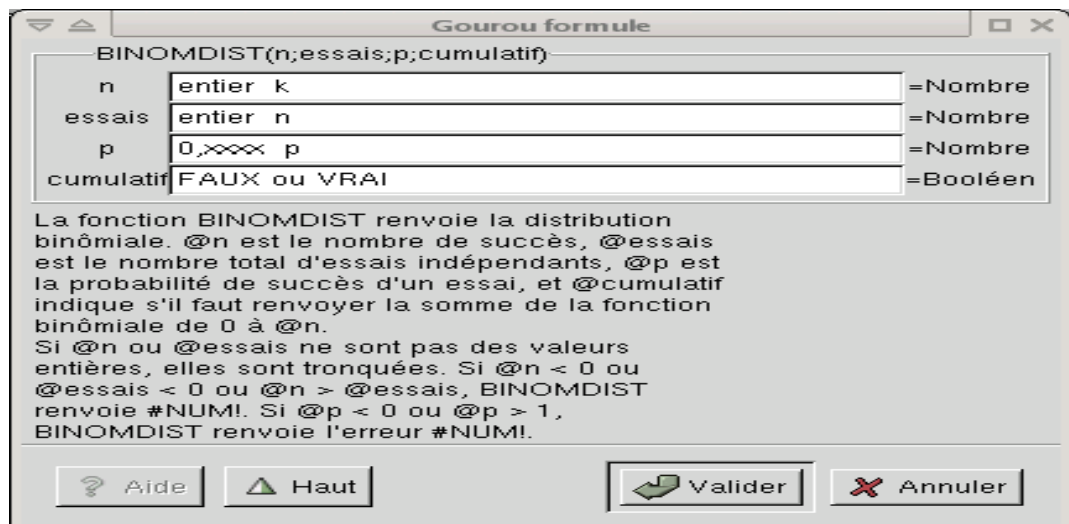


Illustration 1: la fonction binomdist

- (2) Représenter la loi de X à l'aide d'un diagramme bâton
- (3) Quelle est la valeur la plus probable (c'est le mode de la distribution)
- (4) Calculer la probabilité des évènements suivants
 - $[3 \leq X \leq 9]$
 - $[X \leq 10]$
 - $[X \geq 4]$
- (5) Calculer la moyenne et la variance de cette loi
- (6) Calculer et représenter la fonction de répartition. Ici on utilise à nouveau la fonction Binomdist.

Exercice 2 Approximation binomiale-Poisson

- (1) Comparer à l'aide de diagrammes « bâton » la loi de Poisson de paramètre 5 avec les lois binomiales suivantes
 - $B(10, \frac{1}{2})$
 - $B(20, \frac{1}{4})$
 - $B(50, 1/10)$
 - $B(100, 1/20)$
- (2) Pour $n=10$; 20 ; 50 et 100 , évaluer le maximum de l'erreur commise lorsque l'on approxime la loi binomiale par la loi de Poisson, c'est à dire

$$Erreur_n = \max(|P(X_n = k) - P(Y = k)|, k \in \mathbb{N})$$

où X_n suit une loi binomiale de paramètres $\left(n, \frac{5}{n}\right)$ et Y suit une loi de poisson de paramètre 5

(3) Commenter les résultats obtenus.

Exercice 3 Approximation binomiale-Poisson : Application

Suite à une campagne de vaccination contre le paludisme, on estime à 2%, la proportion de personnes qui seront pourtant atteintes de la maladie.

On note X la variable aléatoire égale au nombre de personnes malades dans un petit village de 100 habitants.

- (1) Quelle est la loi de X (préciser les paramètres)?
- (2) Calculer la probabilité de constater plus d'une personne malade ?
- (3) Calculer la probabilité de constater au plus dix personnes malades ?
- (4) Recalculer ces probabilités en utilisant une approximation de la loi de X par une loi de Poisson (préciser le paramètre)
- (5) Quelle est la qualité de l'approximation ?

Exercice 4 : Les sommes partielles

On réalise les opérations suivantes

- La cellule B1 contient la formule $=\text{sum}(\$A\$1:A1)$
- On recopie la cellule B1 dans les cellules B2 à B5.

	A	B	C	D	E
1	1	$=\text{sum}(\$A\$1:A1)$			
2	2				
3	3				
4	4				
5	5				
6					
7					
8					
9					

- On obtient le résultat suivant :

	A	B	C
1	1	1	
2	2	3	
3	3	6	
4	4	10	
5	5	15	
6			
7			

- (1) Expliquer les opérations réalisées et préciser les formules qui apparaissent dans les cellules B2, B3, B4 et B5.
- (2) Plus généralement, si la colonne A contient les valeurs x_1, \dots, x_n , que contient la colonne B après ces opérations?

Exercice 5 : Simulation : la loi binomiale

Un générateur de nombres aléatoires est une fonction qui reproduit les résultats d'une expérience aléatoire. On obtient ainsi une suite de nombres simulés suivant une loi de probabilité donnée.

Par exemple, le générateur de la loi de Bernoulli de paramètre $\frac{1}{2}$ reproduit le lancer d'une pièce équilibrée.

Les générateurs de nombres aléatoires disponibles sous GNUMERIC sont accessibles par le menu

édition/remplir/générateur de nombres aléatoires

- On sélectionne la loi (voir un extrait de la liste Illustration 2),
- On fixe (Voir Illustration 3 et Illustration 4)
 - les paramètres de la loi,
 - le nombre d'échantillons
 - la taille des échantillons

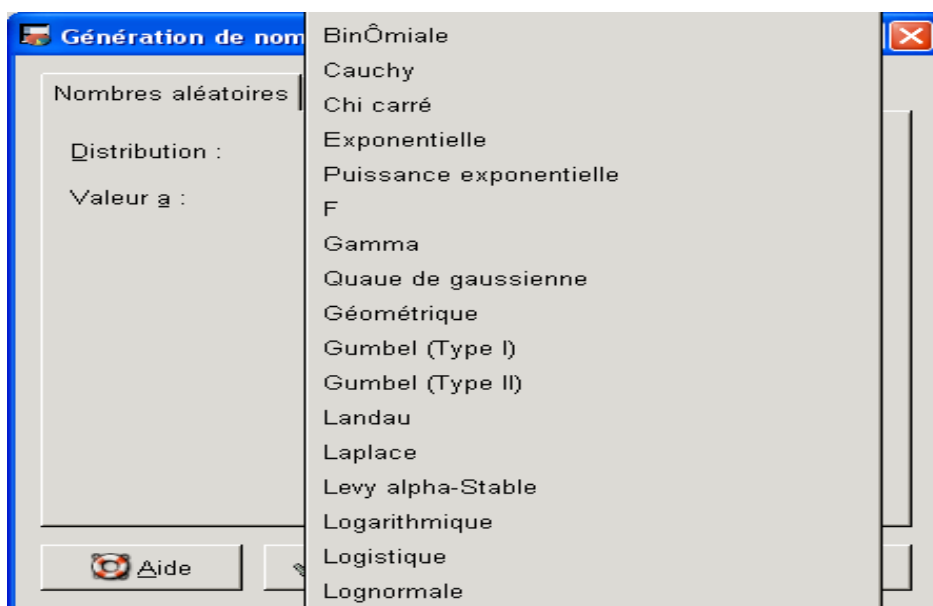


Illustration 2: la liste des générateurs

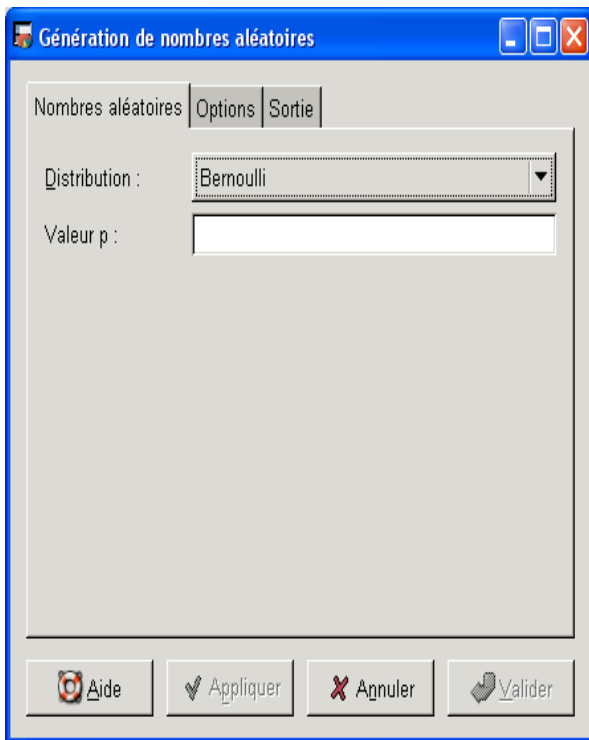


Illustration 3: Choix de la loi et des paramètres

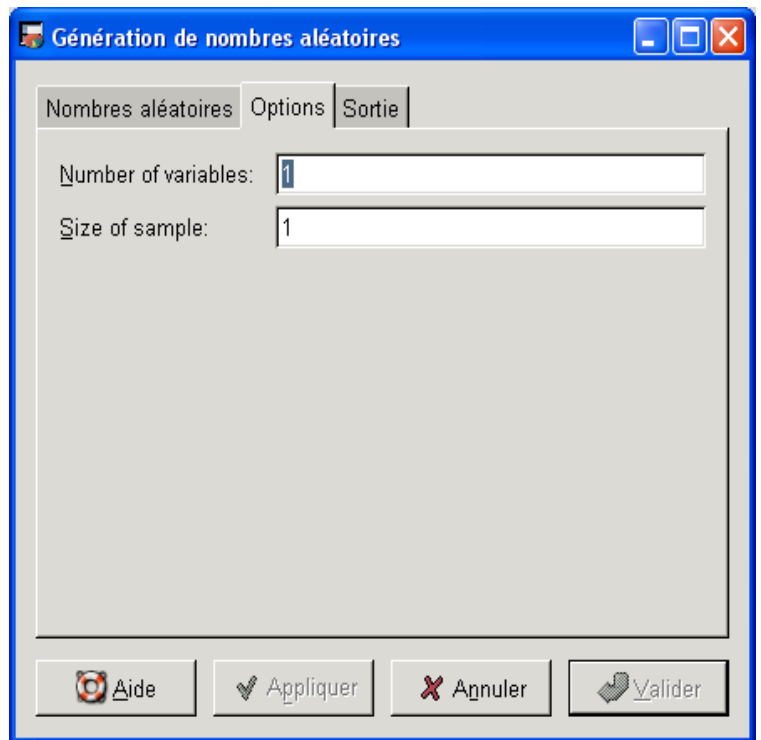


Illustration 4: On fixe le nombre d'échantillons et la taille des échantillons

- (1) Simuler un échantillon de taille 500 distribué suivant la loi binomiale de paramètres $n=15$ et $p=0,3$. On note x_1, \dots, x_{500} les valeurs simulées
- (2) Représenter cet échantillon : tracer le nuage de points $\{(i, x_i) \mid i \in 1, \dots, 500\}$
- (3) Calculer la fréquence des différentes observations $\{0, \dots, 15\}$ en utilisant la fonction frequency. On note f_0, \dots, f_{15} ces fréquences (empiriques).
- (4) Représenter les fréquences f_0, \dots, f_{15} à l'aide d'un diagramme « bâton »
- (5) Comparer la distribution de la loi binomiale de paramètres $(15 ; 0,3)$ et les fréquences empiriques f_0, \dots, f_{15}
- (6) Calculer et représenter graphiquement les fréquences cumulées
Indication : utiliser la méthode décrite dans l'exercice « les sommes partielles ».
- (7) Comparer la fonction de répartition de la loi binomiale de paramètres $(15 ; 0,3)$ et la courbe des fréquences cumulées.
- (8) Calculer la moyenne empirique de l'échantillon simulé.
- (9) Comparer la moyenne empirique et l'espérance de la loi binomiale (voir exercice 1)
- (10) Calculer la variance empirique de l'échantillon simulé.
- (11) Comparer la variance empirique et la variance de la loi binomiale (voir exercice 1)
- (12) Simuler un second échantillon distribué suivant la loi binomiale de paramètres $n=15$ et $p=0,3$ et refaire les questions précédentes.
- (13) Comparer et commenter les résultats obtenus.

Exercice 6 Loi faible des grands nombres

Soit X_1, \dots, X_n n variables aléatoires indépendantes et distribuées suivant la loi de Bernoulli de paramètre p . On veut illustrer la convergence en probabilité de

$$W_n = \frac{X_1 + \dots + X_n}{n} = \frac{S_n}{n} \text{ vers } p.$$

On fixe $p=0,5$ et $n=5$

(1) On commence par simuler un échantillon de taille 100 distribué suivant la loi de W_n . Indication :

→ Quelle est la loi de $S_n = X_1 + \dots + X_n$?

→ Simuler un échantillon de taille $N=100$ distribué suivant la loi de S_n .

On le note s_1, \dots, s_N

→ On déduit un échantillon distribué suivant la loi de W_n en prenant

$$\frac{s_1}{n}, \dots, \frac{s_N}{n}$$

(2) On fixe $\epsilon=0,1$. Évaluer la proportion $R_n(\epsilon)$ de valeurs simulées qui vérifient $|W_n - p| > \epsilon$

(3) Refaire le calcul de $R_n(0,1)$ pour différentes valeurs de $n= 10; 25; 50; 75; 100; 150; 200; 500$

(4) Refaire les calculs pour $\epsilon=0,01$ et $n= 10; 100; 200; 500; 1000; 5000; 10000; 50000$;

(5) Commenter les résultats obtenus