LOI BINOMIALE

Partie 1 : Présentation et traitement en utilisant la loi binomiale sans ordinateur ou calculatrice

Un certain type d'éolienne fonctionne lorsque le vent atteint au moins 8 nœuds. Il faut l'arrêter lorsque le vent dépasse 48 nœuds. Le tableau ci-dessous répertorie les vitesses du vent au cours de l'année :

Vitesse du vent	$[0\,;8[$	$[8\,;24[$	$[24\,;40[$	$[40\,;48[$	$[48\,;55[$
Nbre de jours	70	157	101	24	13

1- Quelle est la fréquence des jours durant lesquels l'éolienne produit de l'électricité ? Arrondir cette valeur au millième.

La vitesse du vent est comprise entre 8 et 48 nœuds 282 jours sur les 365 d'une année. Ainsi l'éolienne produit de l'électricité à une fréquence de $\frac{282}{365} \approx 0{,}773$

On note p cette fréquence. On suppose que p reste constante d'une année sur l'autre et que la vitesse du vent est indépendante d'un jour sur l'autre. On s'intéresse au fonctionnement de l'éolienne au cours d'une semaine. On note X la variable aléatoire qui compte le nombre de jours de fonctionnement de l'éolienne durant la semaine.

2- Justifier que la loi de probabilité de X est une loi binomiale $\mathcal{B}(7\,;\ 0{,}773)$.

On considère ici l'expérience aléatoire suivante : « L'éolienne fonctionne sur une journée tirée au sort ». La probabilité de succès de cette expérience est $p = 0{,}773$.

On répète cette expérience 7 fois afin de simuler ce qui se passe sur une semaine quelconque. D'après l'énoncé, on suppose que les 7 expériences répétées sont indépendantes.

On est ainsi en présence d'une expérience aléatoire au résultat binaire que l'on répète. Cette situation est gérée par un arbre pondéré ou par la loi binomiale dont les paramètres sont ici : $n = 7$ et $p = 0{,}773$. Ainsi La variable X qui donne le nombre de succès sur la série de 7 répétitions suit une loi binomiale $\mathcal{B}(7\,;\ 0{,}773)$

3- Déterminer les coefficients binomiaux qui seront utilisées en complétant le triangle de Pascal ci-dessous :

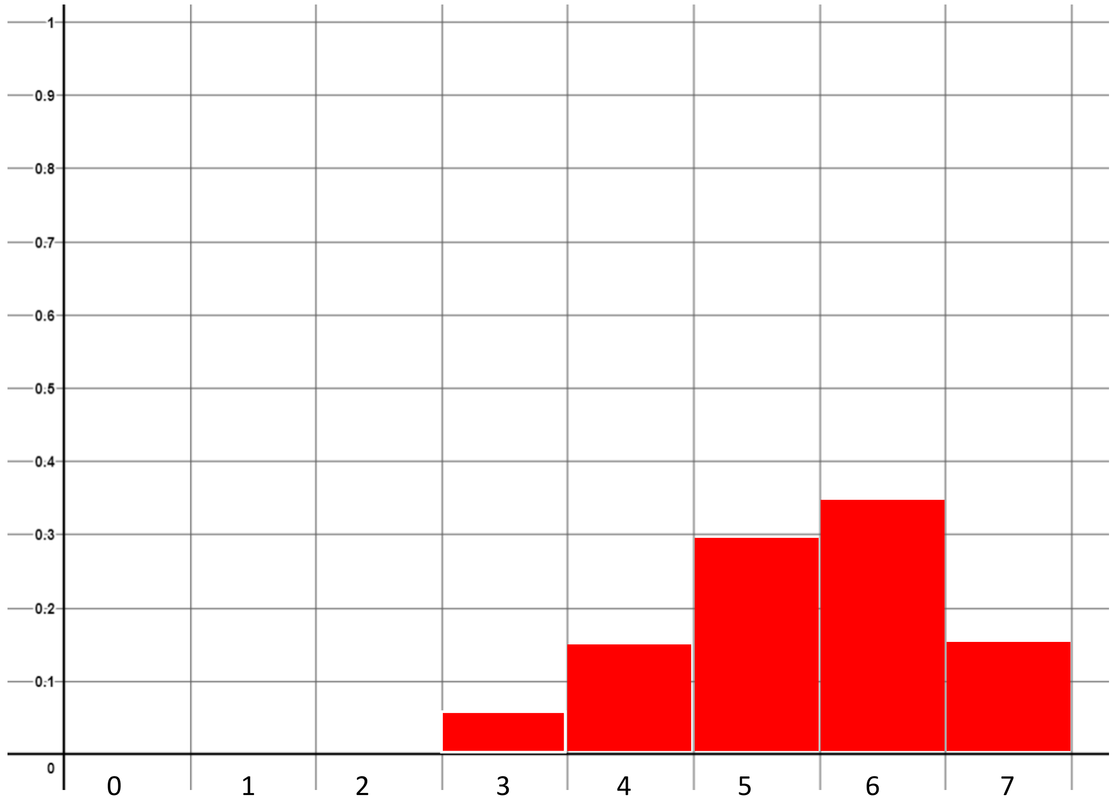
n \ k	0	1	2	3	4	5	6	7
1	1	1						
2	1	2	1					
3	1	3	3	1				
4	1	4	6	4	1			
5	1	5	10	10	5	1		
6	1	6	15	20	15	6	1	
7	**1**	**7**	**21**	**35**	**35**	**21**	**7**	**1**

4- Calculer les probabilités $p(X=0)$, $p(X=1)$, $p(X=2)$, $p(X=3)$, …. d'avoir aucun jour de fonctionnement dans la semaine, 1 seul jour parmi 7 de fonctionnement, 2 jours , …. Arrondir au millième.

Valeurs x_i de X	**0**	**1**	**2**	**3**
$p(X=x_i)$	$\binom{7}{0}0{,}773^0 0{,}227^7$ $1\times 0{,}773^0 0{,}227^7$ 0,000	$\binom{7}{1}0{,}773^1 0{,}227^6$ $7\times 0{,}773^1 0{,}227^6$ 0,001	$\binom{7}{2}0{,}773^2 0{,}227^5$ $21\times 0{,}773^2 0{,}227^5$ 0,008	$\binom{7}{3}0{,}773^3 0{,}227^4$ $35\times 0{,}773^3 0{,}227^4$ 0,043

Valeurs x_i de X	**4**	**5**	**6**	**7**
$p(X=x_i)$	$\binom{7}{4}0{,}773^4 0{,}227^3$ $35\times 0{,}773^4 0{,}227^3$ 0,146	$\binom{7}{5}0{,}773^5 0{,}227^2$ $21\times 0{,}773^5 0{,}227^2$ 0,299	$\binom{7}{6}0{,}773^6 0{,}227^1$ $7\times 0{,}773^6 0{,}227^1$ 0,339	$\binom{7}{7}0{,}773^7 0{,}227^0$ $1\times 0{,}773^7 0{,}227^0$ 0,165

5- Compléter ci-dessous le diagramme donnant ces probabilités sous forme de barres :



<u>**Partie 2**</u> : Traitement en effectuant des simulations aléatoires sur ordinateur avec Excel

Sur ordinateur, lancer Excel.

1- Ouvrir un **nouveau** fichier et l'enregistrer sous le nom *eolienne.xls*

2- Dans la cellule A1, écrire : =ALEA() . Cette fonction d'Excel permet de générer un nombre aléatoire compris entre 0 et 1. A chaque fois que l'on appuie sur la touche F9 du clavier, Excel génère un nouveau nombre …. essayer.

On se sert à présent de cette fonction **alea()** pour pouvoir générer le chiffre 1 si alea() < 0.773 et le chiffre 0 si alea() > 0.773 :

3- Dans la cellule A1, écrire à présent : =SI(ALEA()<0,773;1;0) . Appuyer plusieurs fois sur la touche F9. Le chiffre 1 apparaît dans 77,3 % des cas, sinon c'est le chiffre 0 qui apparaît.

On utilise à présent cette dernière commande pour simuler le fait qu'une éolienne fonctionne sur une journée de la semaine (chiffre 1) ou ne fonctionne pas (chiffre 0).

4- Compléter la ligne 1 en saisissant les titres suivant des colonnes :

	A	B	C	D	E	F	G	H	I
1	**lundi**	**mardi**	**mercredi**	**jeudi**	**vendredi**	**samedi**	**dimanche**		**X**
2									

5- Compléter la ligne 2 en saisissant dans chacune des cellules A2, B2, C2, D2, E2, F2 et G2 la commande (faire du copié-collé): =SI(ALEA()<0,773;1;0)

Saisir dans la cellule I2 la commande : =SOMME(A2:G2) qui permet de compter le nombre de 1 sur la ligne correspondante.

Appuyer plusieurs fois sur la touche F9. Le chiffre de la colonne dont le titre est X donne le nombre de jours de fonctionnement d'une éolienne sur les 7 jours d'une semaine. Par exemple, dans le cas ci-dessous, l'éolienne a fonctionné uniquement 2 jours dans la semaine :

	A	B	C	D	E	F	G	H	I
1	**lundi**	**mardi**	**mercredi**	**jeudi**	**vendredi**	**samedi**	**dimanche**		**X**
2	0	1	0	0	0	1	0		2

On constate qu'en réalisant plusieurs simulations du fonctionnement de l'éolienne sur une seule semaine, les résultats par rapport à la valeur de X sont différents. On peut par-contre déjà constater que les valeurs $X = 5$ ou $X = 6$ reviennent assez souvent. Au contraire la valeur $X = 0$ sort assez rarement. Ainsi il est assez probable que sur une semaine, l'éolienne fonctionne durant 5 ou 6 jours. Il est par-contre très peu probable que l'éolienne ne fonctionne pas du tout sur la semaine.

Pour préciser cela, on se propose de réaliser en une seule exécution, une simulation du fonctionnement de l'éolienne sur 10 semaines différentes.

6- Sélectionner la ligne 2 entre les cellules de A2 à I2. Dupliquer cette ligne 9 fois, en « ***tirant vers le bas*** » le coin bas-droit de cette zone de sélection

	A	B	C	D	E	F	G	H	I
1	**lundi**	**mardi**	**mercredi**	**jeudi**	**vendredi**	**samedi**	**dimanche**		**X**
2	1	1	1	1	1	1	1		7
3									⇩

On obtient ainsi :

	A	B	C	D	E	F	G	H	I
1	**lundi**	**mardi**	**mercredi**	**jeudi**	**vendredi**	**samedi**	**dimanche**		**X**
2	1	1	1	1	1	0	1		6
3	0	1	1	0	1	1	1		5
4	1	1	1	1	1	1	1		7
5	1	1	1	1	1	1	1		7
6	0	1	1	1	1	1	1		6
7	1	1	1	1	1	1	1		7
8	1	1	1	1	1	1	1		7
9	0	0	0	1	1	1	1		4
10	0	1	0	1	1	1	1		5
11	1	1	1	1	1	1	1		7
12									

Pour calculer la fréquence d'apparition de $X = 0$, $X = 1$, …..on utilise la fonction : **NB.SI()** :

7- Reproduisez les colonnes K et L comme sur la figure ci-dessous :

	A	B	C	D	E	F	G	H	I	J	K	L
1	**lundi**	**mardi**	**mercredi**	**jeudi**	**vendredi**	**samedi**	**dimanche**		**X**		**X**	**p(X=xi)**
2	1	1	1	1	1	1	0		6		*0*	
3	1	1	1	0	1	1	1		6		*1*	
4	1	1	1	1	1	0	1		6		*2*	
5	0	1	1	1	1	1	1		6		*3*	
6	0	1	1	1	0	1	1		5		*4*	
7	1	1	1	0	1	1	1		6		*5*	
8	1	1	0	1	1	1	1		6		*6*	
9	1	1	1	1	1	0	1		6		*7*	
10	0	1	1	1	1	1	1		6			
11	0	1	1	1	1	1	1		6			
12												

8- Dans la cellule L2, écrire la commande :

	I	J	K	L
	X		X	
	4		*0*	=NB.SI(I2:I11;K2) NB.SI(**plage**; critère)
	6		*1*	
	6		*2*	
	5		*3*	
	7		*4*	
	6		*5*	
	4		*6*	
	5		*7*	
	6			
	6			

Cette commande permet de compter parmi toutes les cellules entre I2 et I11, celles qui ont leur valeur égale à la valeur de la cellule K2.

9- Dupliquer cette cellule L2 vers le bas afin de répéter cette fonction NB.SI() dans les cellules L3, L4, …. En procédant ainsi, on s'aperçoit que la zone entre I2 et I11 est modifiée. Afin de pouvoir dupliquer correctement la cellule L2 , il faut légèrement modifier la notation utilisée pour définir la zone que l'on ne souhaite pas voir décaler. On remplace pour cela la commande de la cellule L2 par :

=NB.SI(I$2:I$11;K2)
NB.SI(plage; **critère**)

INFO UTILE :

L'écriture de I2:I11 permet de sélectionner les cellules de la colonne I , de la ligne 2 à la ligne 11. Si on duplique cette formule de 2 colonnes vers la droite, cette zone sera modifiée et deviendra K2:K11 .

En écrivant $I$2:$I$11 on sélectionne exactement les mêmes cellules, MAIS dans le cas d'une duplication de cette formule vers la droite ou vers le bas, cette zone ne sera pas modifiée et restera $I$2:$I$11

En écrivant I$2:I$11, le $ n'est placé que devant le nombre des lignes et pas devant la lettre des colonnes. Ainsi, si on duplique cette formule de 2 colonnes vers la droite et de 2 lignes vers le bas, la zone sera modifiée uniquement sur les colonnes et pas sur les lignes. Elle devient K$2:K$11

Finalement on obtient :

	A	B	C	D	E	F	G	H	I	J	K	L
1	**lundi**	**mardi**	**mercredi**	**jeudi**	**vendredi**	**samedi**	**dimanche**		**X**		**X**	**p(X=xi)**
2	1	1	1	0	0	1	1		5		*0*	0
3	1	1	1	1	1	1	1		7		*1*	0
4	0	0	1	1	1	1	0		4		*2*	0
5	1	1	1	1	1	1	1		7		*3*	0
6	1	1	0	1	1	1	0		5		*4*	2
7	1	1	1	1	1	1	1		7		*5*	3
8	1	1	1	1	1	0	1		6		*6*	2
9	1	1	1	1	1	1	0		6		*7*	3
10	1	0	1	1	0	0	1		4			
11	0	0	1	1	1	1	1		5			
12												

I	J	K	L
X	n	X	p(X=xi)
5	10	0	0
4		1	0
7		2	0
4		3	0,1
6		4	0,2
7		5	0,2
3		6	0,3
6		7	0,2
5			
6			

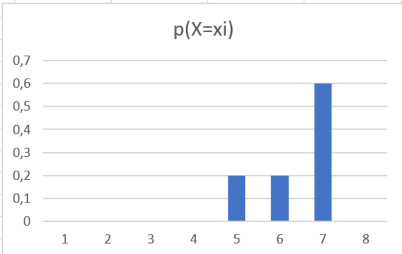
Afin de pouvoir afficher dans le tableau créé des fréquences, ou proportions, on divise les valeurs obtenues avec NB.SI() et qui renvoient un nombre, par le nombre total de semaines simulées : On réexécute l'opération précédente avec cette fois ci dans la cellule L2 la commande :

=NB.SI($I$2:$I$11;K2)/$J$2

10- Refaire plusieurs fois une simulation de 10 semaines en appuyant sur la touche F9.

11- Afin de mieux visualiser les résultats obtenus, on sélectionne la colonne L de L1 à L8 et on insère un histogramme 2d :

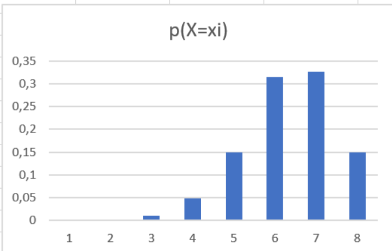
	A	B	C	D	E	F	G	H	I	J	K	L	M
1	lundi	mardi	mercredi	jeudi	vendredi	samedi	dimanche		X	n	X	p(X=xi)	
2	1	1	0	0	1	0	1		4	10	0	0	
3	1	1	0	0	1	1	0		4		1	0	
4	1	1	1	0	1	1	0		5		2	0	
5	1	1	1	1	0	1	1		6		3	0	
6	1	1	0	1	1	1	1		6		4	0,2	
7	1	1	1	1	1	1	0		6		5	0,2	
8	1	1	0	1	1	1	1		6		6	0,6	
9	1	1	1	1	1	0	1		6		7	0	
10	1	1	1	1	0	1	1		6				
11	1	1	0	0	1	1	1		5				



On constate qu'il y a une assez grande dispersion des résultats en relançant la simulation à chaque fois. Afin d'améliorer la précision, on se propose de réaliser cette étude sur cette fois-ci 500 semaines …

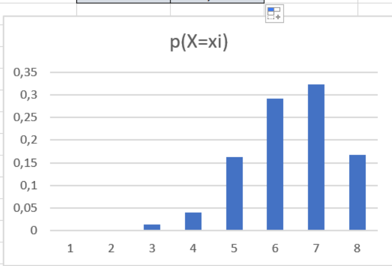
12- Dupliquez à présent la ligne de A2 à I2 jusqu'à la ligne 501. Modifier ce qui doit l'être pour obtenir les mêmes résultats qu'avant, mais sur cette-fois-ci une analyse sur 500 semaines :

1	lundi	mardi	mercredi	jeudi	vendredi	samedi	dimanche		X	n	X	p(X=xi)
2	0	0	1	1	1	1	1		5	500	0	0
3	1	1	0	0	0	1	1		4		1	0,002
4	1	1	1	1	1	1	1		7		2	0,01
5	1	1	1	1	1	1	1		7		3	0,048
6	1	0	0	1	0	1	1		4		4	0,15
7	1	1	1	1	0	1	1		6		5	0,314
8	1	0	1	0	0	1	0		3		6	0,326
9	1	1	0	1	1	1	1		6		7	0,15
10	1	1	0	1	1	1	1		6			
11	1	1	1	0	1	0	1		5			
12	0	1	1	1	1	1	1		6			
13	1	1	0	1	1	1	1		6			
14	1	0	0	1	1	1	0		4			
15	1	1	1	0	1	1	1		6			
16	0	1	1	0	1	1	1		5			
17	1	1	0	0	1	1	0		4			
18	1	1	1	1	1	0	1		6			
19	1	0	1	1	0	1	0		4			
20	1	1	1	0	1	1	1		6			
21	1	1	1	1	1	1	0		6			



13- Recommencez l'opération précédente pour à présent analyser les résultats de 1000 semaines (Dupliquez à présent la ligne de A2 à I2 jusqu'à la ligne 1001). On obtient :

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	lundi	mardi	mercredi	jeudi	vendredi	samedi	dimanche		X	n	X	p(X=xi)	
2	1	1	1	1	1	1	1		7	1000	0	0	
3	1	1	0	0	1	1	1		5		1	0	
4	0	0	1	1	1	0	1		4		2	0,014	
5	1	1	1	1	1	1	0		6		3	0,04	
6	0	0	1	1	1	1	1		5		4	0,163	
7	0	0	1	1	1	1	1		5		5	0,292	
8	1	1	1	1	1	1	0		6		6	0,323	
9	1	1	1	0	1	1	1		6		7	0,168	
10	1	1	0	1	1	1	1		6				
11	1	1	0	1	0	1	1		5				
12	1	1	1	0	1	1	1		6				
13	1	0	1	1	1	1	1		6				
14	1	1	1	1	1	0	1		6				
15	1	0	1	1	1	1	1		6				
16	1	1	0	1	1	1	1		6				
17	1	1	1	1	1	0	0		5				
18	1	1	1	0	0	0	1		4				
19	0	1	1	0	0	0	0		2				
20	1	1	1	1	1	1	1		7				
21	1	1	1	1	1	1	0		6				



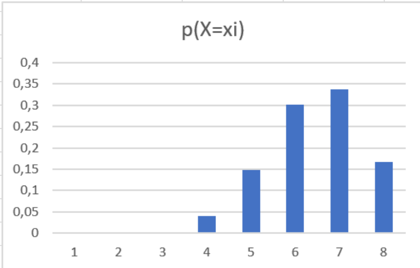
Question : Avec une analyse sur 1000 semaines, la dispersion des résultats est-elle importante lorsque l'on relance une simulation ?

On constate que les résultats varient peu si on réalise notre analyse sur un nombre de semaines important.

Partie 3 : Traitement en utilisant **la fonction loi binomiale** disponible sur Excel

14- Rajouter sur la colonne M , le titre p(X=xi) et différencier cette colonne de la colonne L en utilisant un fond de couleur différent :

	A	B	C	D	E	F	G	H	I	J	K	L	M
1	lundi	mardi	mercredi	jeudi	vendredi	samedi	dimanche		X	n	X	p(X=xi)	p(X=xi)
2	1	1	1	0	0	1	1		5	1000	0	0	
3	1	1	1	1	0	0	0		4		1	0,001	
4	1	1	1	1	1	1	0		6		2	0,003	
5	1	0	0	0	1	1	1		4		3	0,041	
6	0	1	1	1	0	1	0		4		4	0,149	
7	1	1	1	1	1	1	0		6		5	0,302	
8	1	1	1	1	0	1	1		6		6	0,337	
9	1	1	1	1	1	1	1		7		7	0,167	
10	1	0	0	0	0	1	0		2				
11	0	1	0	1	1	1	1		5				
12	1	1	1	1	0	1	1		6				
13	0	0	1	0	1	1	1		4				
14	0	1	0	1	1	1	1		5				
15	1	1	0	1	1	1	0		5				
16	1	1	1	1	1	1	1		7				
17	0	1	0	0	0	1	1		3				
18	0	1	1	1	0	1	1		5				
19	1	1	1	1	1	1	0		6				
20	1	0	1	1	1	1	1		6				
21	1	1	1	1	0	1	1		6				



15- Dans la cellule M2, écrire : =LOI.BINOMIALE.N(K2;7;0,773;) . Dupliquer cette cellule vers le bas jusqu'à la cellule M7. On obtient :

n	X	p(X=xi)	p(X=xi)
1000	0	0	3,10585E-05
	1	0,001	0,000740342
	2	0,003	0,007563235
	3	0,05	0,042924966
	4	0,151	0,146171799
	5	0,297	0,298654098
	6	0,335	0,339000907
	7	0,163	0,164913594

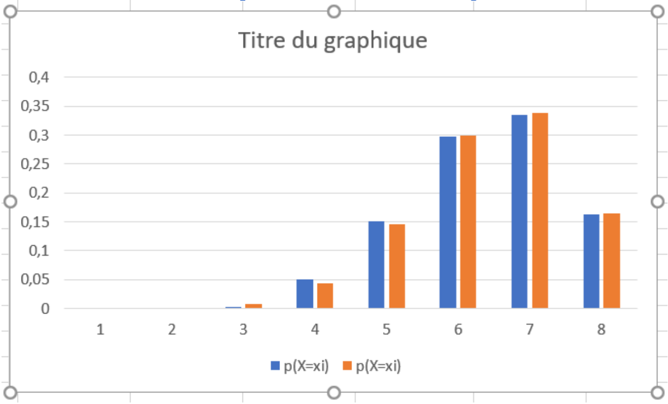
En relançant la simulation, les valeurs de cette colonne M ne changent pas.

Question : Comparer ces valeurs de la colonne M avec celles trouvées en partie 1 dans la question 4 :

On constate que les valeurs calculées par Excel sont identiques à celles déterminées en partie 1.

n	X	p(X=xi)	p(X=xi)
1000	0	0	3,10585E-05
	1	0,001	0,000740342
	2	0,003	0,007563235
	3	0,05	0,042924966
	4	0,151	0,146171799
	5	0,297	0,298654098
	6	0,335	0,339000907
	7	0,163	0,164913594

16- Pour terminer, on modifie l'histogramme tracé avant en ajoutant ces nouvelles valeurs issues de la loi binomiale. On doit obtenir :



Question : La correspondance entre résultats obtenus sur une simulation de 1000 semaines et ceux calculés avec la loi binomiale est-elle correcte ?

La correspondance entre résultats obtenus sur une simulation de 1000 semaines et ceux calculés avec la loi binomiale sont assez proches.

Partie 4 : Espérance et écart-type

17- Avec un click droit sur le J de la colonne J, insérer une colonne

18- Dans la cellule K7, donner le résultat de la moyenne de la colonne des X (colonne I) : =MOYENNE(I2:I1001)

I	J	K	L	M	N
X		n	X	p(X=xi)	p(X=xi)
6		1000	0	0	3,1059E-05
6			1	0,001	0,00074034
6			2	0,003	0,00756323
5			3	0,045	0,04292497
3		X moyen	4	0,135	0,1461718
6			5	0,295	0,2986541
7					

19- Compléter la colonne J afin d'obtenir la valeur des écarts $(X - Xmoyen)^2$: =(I2-$K$7)*(I2-$K$7)

Cela donne :

I	J	K	L	M	N
X	(X-Xmoyen)²	n	X	p(X=xi)	p(X=xi)
6	0,393129	1000	0	0	3,1059E-05
6	0,393129		1	0,001	0,00074034
3	5,631129		2	0,005	0,00756323
7	2,647129		3	0,049	0,04292497
4	1,885129	X moyen	4	0,148	0,1461718
7	2,647129	5,373	5	0,311	0,2986541

20- Dans la cellule K9, donner le résultat de la moyenne de la colonne des $(X - Xmoyen)^2$ (colonne J) :

=RACINE(MOYENNE(J2:J1001))

I	J	K	L	M	N
X	**(X-Xmoyen)²**	**n**	**X**	**p(X=xi)**	**p(X=xi)**
7	2,5921	1000	*0*	0	3,1059E-05
5	0,1521		*1*	0,001	0,00074034
5	0,1521		*2*	0,009	0,00756323
7	2,5921		*3*	0,045	0,04292497
7	2,5921	**X moyen**	*4*	0,154	0,1461718
5	0,1521	5,39	*5*	0,294	0,2986541
5	0,1521	**Ecart-type**			
7	2,5921	1,13132665			

La valeur de la cellule Xmoyen est celle de l'espérance. On la note $E(X)$. Celle de l'écart-type est notée $\sigma(X)$.

21- La théorie mise en place dans la loi Binomiale nous enseigne que pour une loi $\mathcal{B}(n, p)$ l'espérance $E(X)$ et l'écart-type $\sigma(X)$ sont donnés par les formules suivantes : $E(X) = np$ et $\sigma(X) = \sqrt{np(1-p)}$
Utiliser ces formules pour calculer l'espérance et l'écart-type issue de la loi binomiale $\mathcal{B}(7\ ;\ 0{,}773)$.

$E(X) = np = 7 \times 0{,}773 = 5{,}411$ et $\sigma(X) = \sqrt{np(1-p)} = \sqrt{7 \times 0{,}773(1 - 0{,}773)} \approx 1{,}108$

On constate que ces valeurs sont proches de celles trouvées par simulation sur Excel.