Chapitre 7 : Loi binomiale et échantillonnage.

I Attendus.

- Reconnaitre un schéma de Bernoulli et donc une loi binomiale ainsi que déterminer ses paramètres et son espérance.
- Savoir utiliser la machine pour déterminer les probabilité.
- Savoir interpréter l'espérance.
- Savoir déterminer un intervalle de fluctuation au seuil de 95 %.
- Savoir interpréter ce résultat pour une prise de décision.
- Savoir déterminer un intervalle de confiance au seuil de confiance de 95 %.

II Loi binomiale.

A Exemple simple avec arbre pondéré.

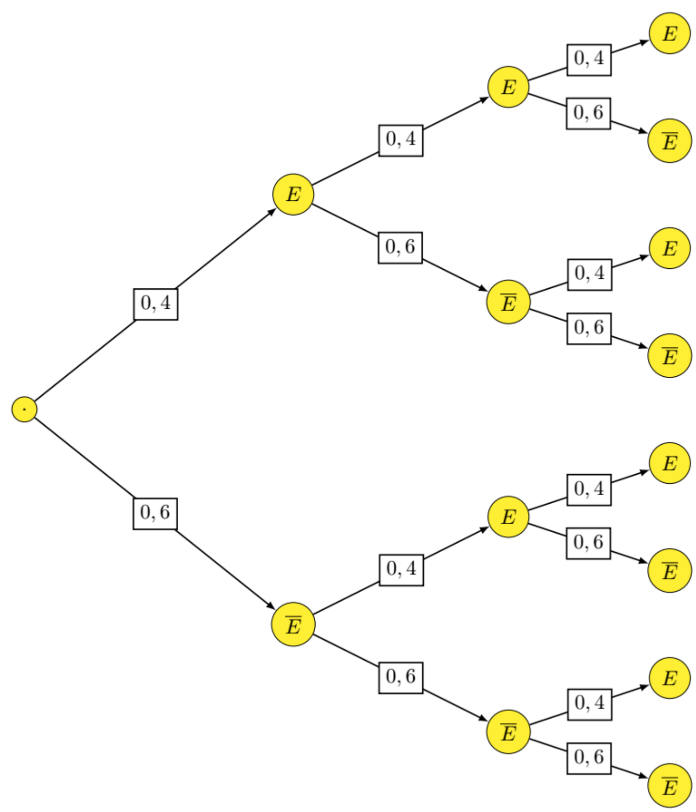
Exemple 1. Dans un établissement, 40 % des élèves sont externes et ne mangent donc pas dans l'établissement pour le repas de midi.

Ce matin, le cuisinier décide d'interroger 3 élèves au portail pour connaitre leurs goûts en matière d'alimentation.

On note E l'événement "l'élève interrogé est externe".

On obtient la loi de probabilité de la variable X donnant le nombre d'élèves interrogés externe parmi les 3 élèves interrogés :

Dans ce cas on dira que la variable X suit la loi binomiale de paramètre 3 et 0,4. On pourra noté simplement : $X \sim \mathscr{B}(3; 0,4)$.



Valeurs possibles : x_i	0	1	2	3
Probabilités $P(X = x_i)$	$0,6^3 = 0,216$	$3 \times 0,4 \times 0,6^2 = 0,432$	$3 \times 0,4^2 \times 0,6 = 0,288$	$0,4^3 = 0,064$

Donc $E(X) = 1,2$.

On a alors simplement $E(X) = np = 3 \times 0,4 = 1,2$.

B Définitions : Schéma de Bernoulli ou loi binomiale.

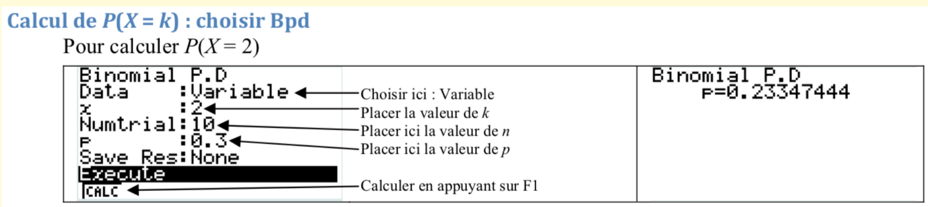
Définition 1

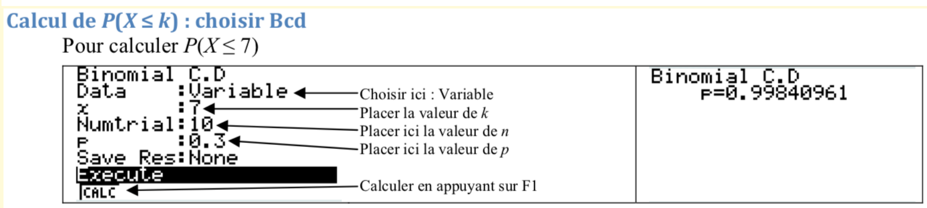
Si l'on réalise n expérience **identiques** avec deux issues possibles. (une telle expérience est appelé schéma de Bernoulli)
On dira de la variable aléatoire X qui donne le nombre de cas où une des issues est réalisée qu'elle suit une loi binomiale de paramètre n et p (p étant la probabilité que l'issue comptabilisé lors d'une des n expériences). On pourra noté $X \sim \mathcal{B}(n;p)$.

C Loi Binomiale.

Méthode 1

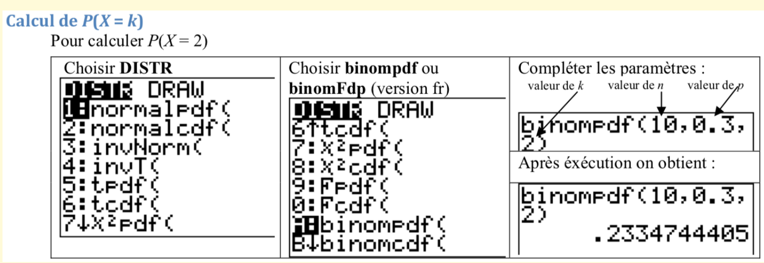
A la calculatrice Casio : Menu "Stat", choisir "DIST", puis "BINM", puis :

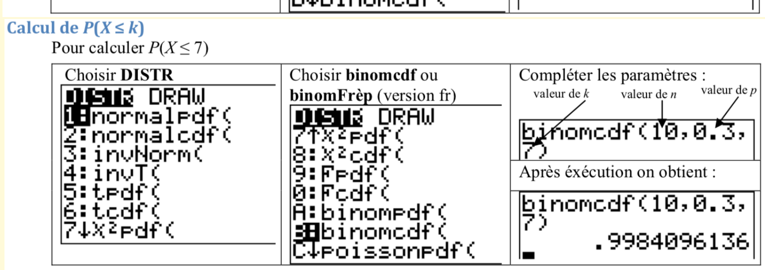




Méthode 2

A la calculatrice TI : Taper "2^{nd} DISTR", puis :





Proposition 1

Si $X \sim \mathcal{B}(n;p)$ alors : $E(X) = np$

III Intervalle de Fluctuation.

Définition 2

On considère une population dont une proportion p des individus possèdent un caractère donné. On prélève dans cette population un **échantillon** de taille n. Soit X la variable aléatoire représentant le nombre d'individus possédant ce caractère dans cet **échantillon**. Alors X suit une loi $\mathscr{B}(n,p)$.

L'intervalle de fluctuation au seuil de 95 % associée à la variable aléatoire X est :

$$\left[\frac{a}{n};\frac{b}{n}\right] \quad \begin{cases} \text{ou est le plus petit entier } a \text{ tel que} P(X \leqslant a) \geqslant 2,5 \\ \text{ou est le plus petit entier } b \text{ tel que} P(X \leqslant a) \geqslant 97,5 \end{cases}$$

Dés lors $P(X \in [a;b]) \simeq 0,95$.

Proposition 2

Dans le cas où

- $n \geqslant 30$ (l'effectif de l'échantillon)
- $np \geqslant 5$ (l'espérance des succès)
- $n(1-p) \geqslant 5$ (l'espérance des échecs)

alors on obtient rapidement un intervalle proche de **l'intervalle de fluctuation au seuil de 95 % associée à la variable aléatoire X** , avec la formule :

$$\left[p - \frac{1}{\sqrt{n}}; p + \frac{1}{\sqrt{n}}\right]$$

Et si l'on veut être plus précis, on utilisera l'intervalle :

$$\left[p - 1,96\sqrt{\frac{p(1-p)}{n}}; p + 1,96\sqrt{\frac{p(1-p)}{n}}\right]$$

IV Prise de décision

Proposition 3

Si l'on fait **l'hypothèse** : "La proportion du caractère dans la population est égale à p" et que la fréquence observé sur **l'échantillon** de taille n est f. Alors :

- Si $f \in \left[\frac{a}{n};\frac{b}{n}\right]$, **l'hypothèse** n'est pas rejetée au seuil de risque d'erreur de 5 %.
- Si $f \notin \left[\frac{a}{n};\frac{b}{n}\right]$, **l'hypothèse** est rejetée au seuil de risque d'erreur de 5 %.

V Estimation d'une proportion.

Proposition 4

On considère une population dont une proportion p des individus possèdent un caractère donné. On prélève dans cette population un **échantillon** de taille n et l'on obtient une fréquence f de cet échantillon possédant le caractère étudié. On appelle **l'intervalle de confiance de la proportion p au niveau de confiance 0,95**, l'intervalle :

$$\left[f-\frac{1}{\sqrt{n}};f+\frac{1}{\sqrt{n}}\right] \qquad \left(\text{ou peu plus précis} \qquad \left[f-1,96\sqrt{\frac{f(1-f)}{n}};f+1,96\sqrt{\frac{f(1-f)}{n}}\right]\right)$$

Il faudra bien sûr $n \geqslant 30$ (l'effectif de l'échantillon), $nf \geqslant 5$ et $n(1-f) \geqslant 5$.