I) Dénombrements classiques.

1) Collections formées de p d'éléments issus d'un ensemble E ayant n éléments.

	avec ordre	sans ordre
avec répétitions possibles	noms : liste de longueur p d'éléments de E ; $p-$liste d'éléments de E ; $p-$uplet d'éléments de E ensemble : E^p nombre : $\lvert E^p\rvert = ...$	noms : combinaison avec répétitions d'ordre p d'éléments de E ; $p-$combinaison avec répétition d'éléments de E ensemble : nombre : hors programme
sans répétition	noms : liste de longueur p d'éléments distincts de E ; $p-$liste d'éléments distincts de E ; $p-$uplet d'éléments distincts de E ; $p-$ arrangement d'éléments de E ensemble : $\mathcal{L}_p(E)$ nombre : $\lvert\mathcal{L}_p(E)\rvert = A_n^p = ... =$	noms : sous-ensemble de taille p de E ; $p-$sous-ensemble de E ; partie de taille p de E ; $p-$partie de E ; combinaison d'ordre p d'éléments de ; $p-$combinaison d'éléments de E ensemble : $\mathcal{P}_p(E)$ (ou $\binom{E}{p}$) nombre : $\lvert\mathcal{P}_p(E)\rvert = ... = \frac{.................}{.................}$

On peut considérer que les n éléments de E sont rangés dans une urne et que l'on effectue un tirage de ces éléments ; on a alors le tableau superposable au précédent :

tirage	successif	simultané
avec remise	nom : liste ensemble : E^p nombre : $\lvert E^p\rvert = ...$	Pas de sens
sans remise	nom : arrangement ensemble : $\mathcal{A}_p(E)$ nombre : $\lvert\mathcal{A}_p(E)\rvert = ...$	nom : combinaison ensemble : $\mathcal{P}_p(E)$ nombre : $\lvert\mathcal{P}_p(E)\rvert = ...$

REM : le tirage simultané peut être effectué de manière successive ... L'important étant que l'on ne tient pas compte de l'ordre dans lequel les tirages ont été effectués.

Démonstration de ces divers dénombrements : D5

Rappel : n^p est aussi le nombre d'applications d'un ensemble à .... éléments vers un ensemble à ....éléments
et A_n^p est aussi le nombre d'injections d'un ensemble à .... éléments vers un ensemble à ....éléments.

2) Listes à deux objets (pile ou face).

PROP 1 : le nombre de listes de taille n faisant intervenir deux objets au plus (= le nombre de résultats d'une succession de n jeux de pile ou face) vaut

......

PROP 2 : le nombre de listes de taille n faisant intervenir p fois un objet et $n-p$ fois un autre (= le nombre de résultats d'une succession de n jeux de pile ou face avec p piles) vaut

......

D6

3) Résumé des propriétés des coefficients binomiaux.

Diverses définitions équivalentes.

DEF 1. Définition par récurrence de Pascal	a. $\forall n \in \mathbb{N}$ $\binom{n}{0} = \binom{n}{n} = 1$ b. $\forall n \in \mathbb{N}^*$ $\forall p \in [\|1, n\|]$ $\binom{n+1}{p} =$ ....
DEF 2. Définition par récurrence du pion	a. $\forall n \in \mathbb{N}$ $\binom{n}{0} = 1$ b. $\forall n \in \mathbb{N}^*$ $\forall p \in [\|1, n\|]$ $\binom{n}{p} = \frac{n}{p}\binom{n-1}{p-1}$.
DEF 3. Définition binomiale	$\binom{n}{p}$ est le coefficient de $a^{n-p}b^p$ dans $(a+b)^n$
DEF 4. Définition calculatoire	$\binom{n}{p} = \frac{............}{}$
DEF 5. Définition combinatoire 1	$\binom{n}{p}$ est le nombre de ....
DEF 6. Définition combinatoire 2	$\binom{n}{p}$ est le nombre de

Nous sommes partis de la DEF 1 et avons démontré les caractérisations 2, 3, 4, 5.

Montrons que la DEF 5 entraîne la 3 (démonstration combinatoire de la formule du binôme).
D7

Redémontrons de manière combinatoire que $\sum_{k=0}^{n} \binom{n}{k} = 2^n$, puis que $\sum_{0 \leqslant 2k \leqslant n} \binom{n}{2k} = \sum_{0 \leqslant 2k+1 \leqslant n} \binom{n}{2k+1} = 2^{n-1}$.
D8

Montrons que la DEF 5 entraîne la 1 (démonstration combinatoire de la relation de Pascal).

D9

II) Probabilités sur un ensemble fini.
1) Univers et évènements.

DEF : l'ensemble des résultats d'une expérience aléatoire finie est dénommé "univers (des possibles)", noté traditionnellement Ω, et ses parties s'appellent des "évènements".

Les évènements sont en général définis à partir de la propriété qui les caractérise, avec la notation :

$$\{P\} = \{\omega \in \Omega \ / \ \omega \text{ a la propriété } P\}$$

Par exemple, si l'expérience aléatoire consiste en un lancer de deux dés, alors

$$\begin{aligned} \Omega &= \text{............} \\ \{\text{total des points} = 7\} &= \{\text{...}\} \end{aligned}$$

On a évidemment

$$\begin{aligned} \{P \text{ ou } Q\} &= \{P\} \cup \{Q\} \\ \{P \text{ et } Q\} &= \{P\} \cap \{Q\} \\ P \text{ implique } Q &\Leftrightarrow \{P\} \subset \{Q\} \end{aligned}$$

ATTENTION : mathématiquement, l'évènement, qui est un ensemble, est différent de la propriété qui le détermine, mais, en probabilités, on fait souvent la confusion ; on dira par exemple : "l'évènement A est réalisé" alors que l'on ne réalise pas un ensemble, mais une propriété !

Vocabulaire :

Les singletons de Ω sont les évènements *élémentaires*,

$\varnothing$ l'évènement *impossible*,

Ω l'évènement *certain* ;

le *contraire* d'un évènement est son complémentaire d'un point de vue ensembliste ;

deux évènements sont dits *incompatibles* si leur intersection est vide.

2) Probabilités.

DEF : une probabilité P sur un univers fini Ω est une application de $\mathcal{P}(\Omega)$ dans $\mathbb{R}_+$ vérifiant :

$$\begin{aligned} P(\Omega) &= 1 \\ \text{Si } A \text{ et } B \text{ sont des évènements incompatibles, } P(A \cup B) &= P(A) + P(B) \end{aligned}$$

PROP : premières propriétés :

$$\begin{aligned} P_1 &: \quad P(\varnothing) = 0 \\ P_2 &: \quad P(\overline{A}) = 1 - P(A) \\ P_3 &: \quad A \subset B \Longrightarrow P(A) \leqslant P(B) \\ P_4 &: \quad P(A) \leqslant 1 \\ P_5 &: \quad P(A) + P(B) = P(A \cup B) + P(A \cap B) \ \text{ : formule de Poincaré} \\ P_6 &: \quad \text{si } (A_i)_{i \in I} \text{ est une famille finie d'évènements deux à deux incompatibles :} \quad (1) \end{aligned}$$

$$P\left(\bigcup_{i \in I} A_i\right) = \sum_{i \in I} P(A_i) \quad (2)$$

D10

REM 1 : la probabilité d'un événement A doit être vue comme la limite de la fréquence du cas où A est réalisé lorsque l'on répète un grand nombre de fois l'expérience aléatoire. Une probabilité de 0,6 signifie que si on répète par exemple 100 fois l'expérience, il y aura environ 60 cas où A est réalisé, et 40 où il ne l'est pas.

"$P(A) = 0,6$" sera donc souvent énoncé sous la forme : il y a 60% de chances que A soit réalisé.

Autre façon de dire la même chose : si on répète l'expérience, le fait qu'un évènement ait une probabilité p signifie que l'évènement arrivera en moyenne tous les $1/p$ coups (par exemple : tous les 25 coups pour une proba de 0,04). L'inverse de la probabilité est la périodicité en moyenne de l'évènement.

Inversement, un test statistique permet de définir une probabilité : si sur 1000 patients atteints d'une certaine maladie bien choisis, un médicament X en a sauvé 732, on pourra poser une probabilité de 0,7 de guérison pour X.

REM 2 : on connaît entièrement une probabilité si on la connaît sur les évènements élémentaires.

En effet

$$P(A) = \sum_{\omega \in A} P(\{\omega\})$$

Toute probabilité peut donc être définie par la donnée d'une application p de Ω dans $\mathbb{R}_+$ en posant

$$P(A) = \sum_{\omega \in A} p(\omega)$$

à condition que p vérifie la condition de normalisation

$$\sum_{\omega \in \Omega} p(\omega) = 1$$

Exemple simpliste : si dans un questionnaire les réponses sont "oui" à 60%, "non" à 30% et "ne sait pas" à 10%, on prendra $\Omega = \{\text{oui, non, ne sait pas}\}$, $p(\text{oui}) = 0,6$; $p(\text{non}) = 0,3$; $p(\text{ne sait pas}) = 0,1$; alors, $P(\{\text{oui,non}\}) = 0,6 + 0,3 = 0,9$ etc...

REM 3 : il est possible qu'un événement non vide soit de probabilité nulle (cela arrive plus souvent dans le cas de probabilités sur des univers infini). Un tel évènement est dit "presque (ou quasi-) impossible" ou "négligeable" ; de même, un évènement de probabilité 1 est dit "presque certain" ; et si $P(A \cap B) = 0$, A et B sont dits "presque incompatibles".

PROP et DEF : l'*équiprobabilité* (ou probabilité uniforme) est l'unique probabilité où tous les évènements élémentaires ont même probabilité.
Cette probabilité est définie par :

$$P(A) = \frac{|A|}{|\Omega|} = \frac{\text{nombre de cas favorables (à l'évènement } A)}{\text{nombre de cas total}}$$

D11
REM : on reconnaît l'équiprobabilité lorsque dans l'énoncé on trouve l'expression "au hasard", ou "dés non pipés" etc...

DEF : un système (presque) complet d'évènements est une liste $(A_i)_{i=1..n}$ d'évènements deux à deux (presque) incompatibles dont la réunion est (presque) certaine.
Exemple : $(A, \overline{A})$.

L'intéret de cette notion vient de ce que si $(A_i)_{i=1..n}$ est une telle famille, alors

$$P(A) = \sum_{i=1}^{n} P(A \cap A_i)$$

(formule de filtration ou "du coupe-patate")
D12

III) CONDITIONNEMENT ET INDÉPENDANCE
1) Probabilités conditionnelles.

DEF : si A et B sont deux évènements, B de probabilité non nulle , on pose

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

$P(A|B)$ est lu "probabilité que A soit réalisé, sachant que B l'est", ou, en simplifié "P de A sachant B".

REM 1 : l'idée est rapporter la probabilité de $A \cap B$ à celle de B de sorte que $P(B|B) = 1$.

REM 2 : on a donc, si A et B sont de proba non nulle (petite formule des probabilités composées)

$$P(A \cap B) = P(A|B) P(B) = P(B|A) P(A)$$

ainsi que (petite formule de Bayes)

$$P(A|B) = P(B|A) \frac{P(A)}{P(B)}$$

REM 3 : on a évidemment : A et B sont presque incompatibles ssi $P(A|B) = 0$, ssi $P(B|A) = 0$.

PROP : si $A_1, A_2, ..., A_n$ sont n évènements d'intersection de probabilité non nulle, on a la formule dite "des probabilités composées" :

$$P(A_1 \cap A_2 \cap ... \cap A_n) = P(A_1) P(A_2|A_1) P(A_3|A_1 \cap A_2) ... P(A_n|A_1 \cap A_2 \cap ... \cap A_{n-1})$$

D13
Exemple d'application :
Un urne contient n boules dont b blanches ; on tire sans remise r boules. Quelle est la probabilité que les r boules soient blanches ?
Soit $B_k = \{$la k-ième boule est blanche$\}$; on cherche $P(B_1 \cap B_2 \cap ... \cap B_r) = p$.

En écrivant $p = P(B_1) P(B_2|B_1) P(B_3|B_1 \cap B_2) ... P(B_n|B_1 \cap B_2 \cap ... \cap B_{n-1})$, on obtient

$$p = \frac{b}{n} . \frac{b-1}{n-1} .... \frac{b-r+1}{n-r+1}$$

PROP : si $(A_i)_{i=1..n}$ est un système quasi-complet d'évènements, et B un évènement donné, alors

$$P(B) = \sum_{i=1}^{n} P(B|A_i) P(A_i) \text{ (formule des probabilités totales)}$$

avec la convention que si $P(A_i)$ est nul, alors $P(B|A_i) P(A_i)$ est nul.

$$\text{si } P(B) > 0, P(A_j|B) = \frac{P(B|A_j) P(A_j)}{\sum_{i=1}^{n} P(B|A_i) P(A_i)} \text{(grande formule de Bayes, ou formule de la probabilté des causes)}$$

D14

Exemple : Une population animale comporte $p = 1/3$ de mâles. L'albinisme frappe $q = 6\%$ des mâles et $r = 0,36\%$ des femelles.

Quelle est la proportion s d'albinos dans la population ?

Quelle est la proportion t des mâles parmi les albinos, et la proportion u des mâles parmi les non albinos ?

REP : $s = pq + (1-p) r = ...\%$, $t = \dfrac{pq}{pq + (1-p) r} = ...\%$, $u = \dfrac{p(1-q)}{p(1-q) + (1-p)(1-r)} = ...\%$.

PROP : l'application P_B de $\mathcal{P}(\Omega)$ dans $\mathbb{R}_+$ définie par $P_B(A) = P(A|B)$ est une probabilité sur Ω appelée *probabilité conditionnelle relative* à B.

D15

REM : pour P_B, tous les évènements contenant B sont presque certains !

2) Évènements indépendants.

TH et DEF : si A et B sont deux évènements de probabilité non nulle alors

$$P(A|B) = P(A) \text{ équivaut à } P(B|A) = P(B)$$

ces deux conditions étant équivalentes à $P(A \cap B) = P(A) P(B)$.

On dit que dans ce cas, les événements sont *indépendants* (et cette définition est étendue aux cas où $P(A)$ ou $P(B)$ est nul).

D16

PROP : si A et B sont indépendants, A et $\overline{B}$ également, donc aussi $\overline{A}$ et $\overline{B}$.

D17

Attention : deux évènements peuvent être indépendants , et ne plus l'être sachant un autre !

D18 : A et B sont indépendants sachant C ssi $P(A \cap C) P(B \cap C) = P(A \cap B \cap C) P(C)$; un cas où $(A \cap B \cap C)$ est vide fournit un contre exemple.

IV) VARIABLES ALÉATOIRES.

1) Définition.

DEF : Une variable aléatoire (sous-entendu : réelle) sur un univers Ω est une application de Ω dans $\mathbb{R}$.

Dans ce cours, Ω étant fini, une variable aléatoire prend un nombre fini de valeurs $x_1, ..., x_p$.

Exemple important : si A est un évènement, X_A est la variable aléatoire qui vaut 1 si l'évènement A est réalisé, 0, sinon, (X_A est en fait la fonction caractéristique de la partie A), Concrètement, c'est mon gain si je gagne 1 euro si A se réalise, 0 euro sinon.

On note $\{X = x_i\}$ l'évènement $X^{-1}(x_i) = \{\omega \in \Omega \ / \ X(\omega) = x_i\}$.
La famille des $\{X = x_i\}$ forme donc un système complet d'évènements liés à la variable aléatoire X.
Note : on simplifie l'écriture de $P(\{X = x_i\})$ en $P(X = x_i)$.

DEF : l'application p_X définie sur l'univers image $X(\Omega) = \{x_1, ..., x_p\}$ qui à x_i fait correspondre $P(X = x_i)$ s'appelle la *loi de probabilité* de X. Et on définit sur $X(\Omega)$ la *probabilité induite par* X : P_X par

$$P_X(U) = P(X \in U) \left(= P\left(X^{-1}(U)\right)\right)$$

Exemple : Épreuve aléatoire : lancer de 2 dés.
Univers : $\Omega = [|1, 6|]^2$, avec équiprobabilité.
Variable aléatoire : X = somme des nombres indiqués sur les dés ($X((i, j)) = i + j$).
Univers image : $X(\Omega) = [|2, 12|]$
Evènements élémentaires liés à X : $\{X = k\} = \{(i, j) / i + j = k, 1 \leq i, j \leq 6\}$
Loi de probabilité de X : $P(X = k) = \dfrac{|\{X = k\}|}{36}$ que l'on peut présenter dans le tableau :

2	3	4	5	6	7	8	9	10	11	12
1/36	2/36	3/36	4/36	5/36	6/36	5/36	4/36	3/36	2/36	1/36

Conseil : si vous n'arrivez pas à calculer $P(X = x_i)$, essayez de calculer $P(X \geqslant x_i)$ ou $P(X \leqslant x_i)$; si les x_i sont classés dans l'ordre croissant vous pourrez utliser

$$P(X = x_i) = P(X \geqslant x_i) - P(X \geqslant x_{i+1}) = P(X \leqslant x_i) - P(X \leqslant x_{i-1})$$

2) Espérance, variance, écart-type.

DEF : l'*espérance* d'une variable aléatoire est la moyenne des valeurs prises par cette variable, pondérée par les probabilités correspondantes :

$$E(X) = \sum_{\omega \in \Omega} P(\{\omega\}) X(\omega) = \sum_{k \in X(\Omega)} P(X = k) k$$

Exemples :
- une variable aléatoire constante a cette constante pour espérance.
- dans l'exemple des dés précédent $E(X) = ...$
- Si A est un évènement :

$$E(X_A) = ...$$

Autrement dit : la probabilité d'un évènement, c'est l'espérance de la fréquence de réalisation de cet évènement.

P1 : l'espérance est toujours comprise entre la plus basse et la plus haute valeur prise par la variable.
P2 : théorème de transfert.
Si f est une fonction réelle définie sur $X(\Omega)$

$$E(f(X)) = \sum_{k \in X(\Omega)} P(X = k) f(k)$$

(démontré seulement pour f injective).
P3 : l'espérance est "linéaire", ce qui signifie que

$$E(\alpha X + \beta Y) = \alpha E(X) + \beta E(Y)$$

P4 : elle est croissante :

$$X \leqslant Y \Rightarrow E(X) \leqslant E(Y)$$

D19

Remarque utile en pratique : il arrive assez souvent que $X(\Omega) = [|1, n|]$ (ou $[|0, n|]$), dans ce cas, on a:

$$E(X) = P(X \geqslant 1) + P(X \geqslant 2) + ... + P(X \geqslant n)$$

D20

Application : espérance du nombre d'évènements réalisés parmi n évènements.

PROP : Etant donnés n évènements $A_1, ..., A_n$ (indépendants ou non) , l'espérance du nombre d'évènements réalisés parmi ces n évènements est la somme des probabilités de chacun de ces évènements.

D21

Exemple ; si n évènements sont de probabilité $1/2$, le nombre d'évènements réalisés en moyenne vaut : ....

DEF : la *variance* d'une variable aléatoire est l'espérance du carré des écarts à la moyenne de cette variable ; l'écart-type (standard deviation en anglais) est la racine carrée de la variance :

$$\begin{aligned} V(X) &= E\left((X - E(X))^2\right) \\ \sigma(X) &= \sqrt{V(X)} \end{aligned}$$

REM : pour mesurer les écarts à la moyenne, il pourrait sembler plus naturel de calculer l'*écart moyen* $E(|X - E(X)|)$, mais celui-ci est souvent bien plus difficile à calculer.

PROP : la variance se calcule plus simplement par la formule (de Koenig-Huygens) :

$$V(X) = E\left(X^2\right) - (E(X))^2$$

D22

Exemples :

- la variance et l'écart type d'une variable constante sont nuls.
- dans l'exemple précédent $E\left(X^2\right) =$ .... donc $V(X) =$ ...... et $\sigma(X) =$ ....
- Si X_A est la fonction caractéristique de l'évènement A,

$$V(X_A) = ...$$

PROP :

$$V(aX + b) = a^2 V(X)$$

D23

DEF : une variable aléatoire est dit *centrée* si son espérance vaut 0, et *réduite* si sa variance vaut 1.

PROP : si X est de variance non nulle, $\dfrac{X - E(X)}{\sigma(X)}$ est centrée réduite.

D24