Salle : BIOSTATDUFEU

VARIABLE ALEATOIRE, LOI DE PROBABILITE DISCRETE ET CONTINUE

BIOSTATISTIQUES / UE4 – Pr Staccini

PACES

PLAN :

I- DEFINITIONS

II- VARIABLE ALEATOIRE DISCRETE

III- LOI DE PROBABILITE DISCRETE

IV- VARIABLE ALEATOIRE CONTINUE

V- LOI DE PROBABILITE CONTINUE

I- DEFINITIONS

Une variable aléatoire est une épreuve menant à des évènement élémentaires qui sont des nombres.

- **Discrète** si le résultat fait partie d'un ensemble fini ou dénombrable.
- **Continue** si le résultat est compris dans $\mathbb{R}$ ou un intervalle de $\mathbb{R}$. On l'appelle aussi variable à densité.

II- VARIABLE ALEATOIRE DISCRETE

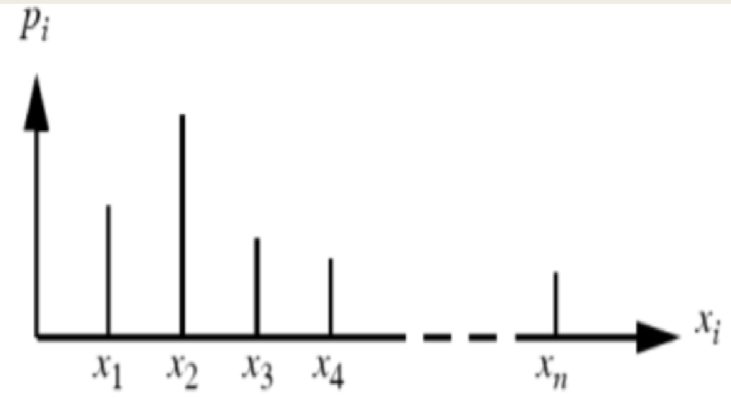
- La loi de probabilité d'une variable aléatoire X discrète est définie en donnant l'ensemble des valeurs $p1, p2, \cdots , pn$ qui sont les probabilités de ses différentes éventualités $x1, x2 \ldots , xn$.

  Soit $pi = P(X = xi\ )$

  donc $0 \leq pi \leq 1\ et \sum pi = 1$

- On peut représenter cette loi par une table ou un diagramme en bâtons.

x_1	x_2	$\ldots$	x_n	$\ldots$
p_1	p_2	$\ldots$	p_n	$\ldots$



A. MOYENNE **μ** :

La moyenne **μ** de la variable aléatoire X est la valeur moyenne des résultats que l'on obtiendrait en répétant indéfiniment l'épreuve.

C'est un indicateur de **position**.

$$\mu = x_1p_1 + x_2p_2 + \cdots + x_np_n = \sum_{i=1}^{i=n} x_ip_i$$

B. ESPERANCE E(X) :

L'espérance est synonyme de la moyenne en probabilités et statistiques.

Théorèmes de l'espérance :

- *Soit X une v. aléatoire et k une constante réelle :* $E(kX) = k\,E(X)$

  $$E(k + X) = E(X) + k$$

- *Soit X et Y deux variables aléatoires définies sur le même espace fondamental :*
  $E(X+Y) = E(X) + E(Y)$

C'est un indicateur de **position,** traduisant la tendance centrale de la v.a.

C. VARIANCE ET ECART-TYPE σ^2 et σ :

- Notée $\boldsymbol{\sigma^2}$, la variance est la moyenne des carrés des écarts à la moyenne.

$$\sigma^2 = Var(X) = \sum_{i=1}^{i=n} p_i (x_i - \mu)^2$$

- L'écart-type est sa racine carrée, $\boldsymbol{\sigma}$.

$$\sigma^2 = Var(X) = E(X - \mu)^2 = E(X^2) - E(X)^2 = E(X^2) - \mu^2$$

Ce sont des indicateurs de **dispersion**.

D. VARIABLE CENTREE REDUITE :

- Soit X une v.a de moyenne $\boldsymbol{\mu}$ et d'écart type $\boldsymbol{\sigma}$, on définit la variable:

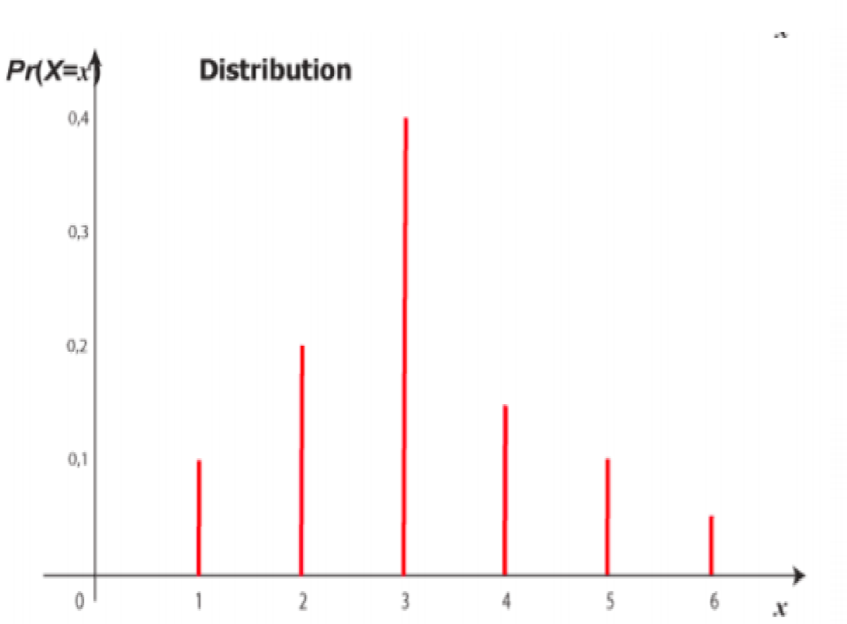
$$Y = \frac{X - \mu}{\sigma}$$

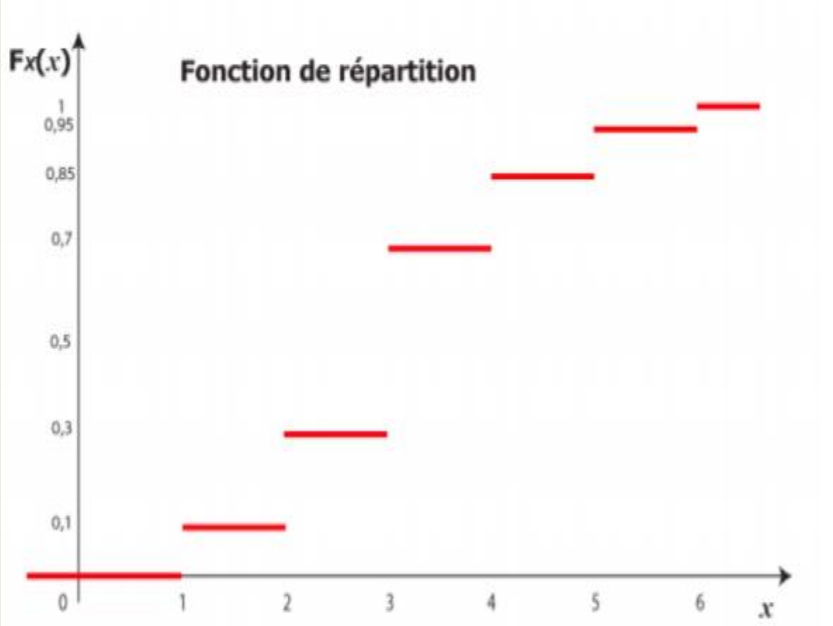
- Y est une variable centrée réduite avec

E(Y) = 0 et Var (Y) = 1

E. FONCTION DE REPARTITION:

- On la définit comme ***F(x) = P(X≤x).***
- C'est une fonction **cumulative** car on additionne toutes les probabilités (pi) des xi survenus avant x.
- Elle est toujours **monotone croissante,** avec : $\lim_{x \to -\infty} F(x) = 0$ et $\lim_{x \to +\infty} F(x) = 1$





QRU SOCRATIVE :

Salle : BIOSTATDUFEU

À propos des variables aléatoires, donnez la réponse vraie:

A- L’écart-type est un indicateur de position

B- L’écart-type est la racine carrée de la variance

C- Une variable centrée réduite est définit par Y = $\frac{X-\mu}{\sigma}$ avec E(Y) = Var(Y) = 1

D- L’Espérance en statistiques est synonyme d’écart-type (µ)

E- Les réponses ABCD sont fausses.

CORRECTION : **REPONSE B**

À propos des variables aléatoires, donnez LA réponse vraie :

A- L'écart-type est un indicateur de ~~position~~ **dispersion**

B- L'écart-type est la racine carrée de la variance

C- Une variable centrée réduite est définit par Y = $\frac{X-\mu}{\sigma}$ avec ~~E(Y) =~~ Var(Y) = 1 et **E(Y) = 0**

D- L'Espérance en statistiques est synonyme ~~d'écart-type~~ (µ) de **moyenne**

E- les réponses ABCD sont fausses.

III- LOIS DE PROBABILITE DISCRETE

A- loi de BERNOULLI B(p)

L’Epreuve de Bernoulli est une expérience aléatoire dont l’issue est un « succès » ou un « échec ».

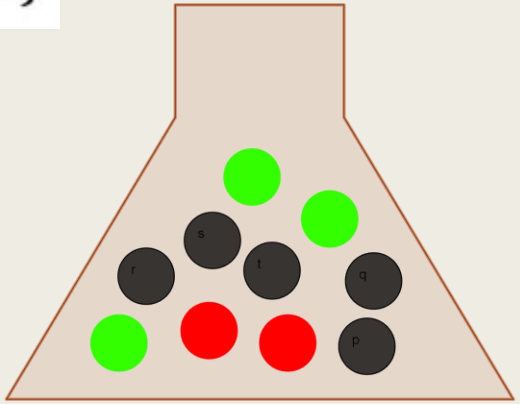
- Paramètres : p : probabilité d’un succès

  q = 1-p : probabilité de l’échec

  X est la v.a donnant le nombre de « succès » pendant l’épreuve (0 ou 1).

$$P(X=k) = p^k(1-p)^{1-k} = p^k q^{1-k} \qquad \text{pour} \qquad k \in \{0,1\}$$

- **μ=p**
- **σ^2 = p(1-p) = pq**



Rappel : $C_n^k = \frac{n!}{k!(n-k)!}$

B- loi BINOMIALE B(n ; p)

La loi **Binomiale** consiste en la répétition de n épreuves de Bernoulli indépendantes.

❖ Paramètres : **n** : nombre d'essais indépendants

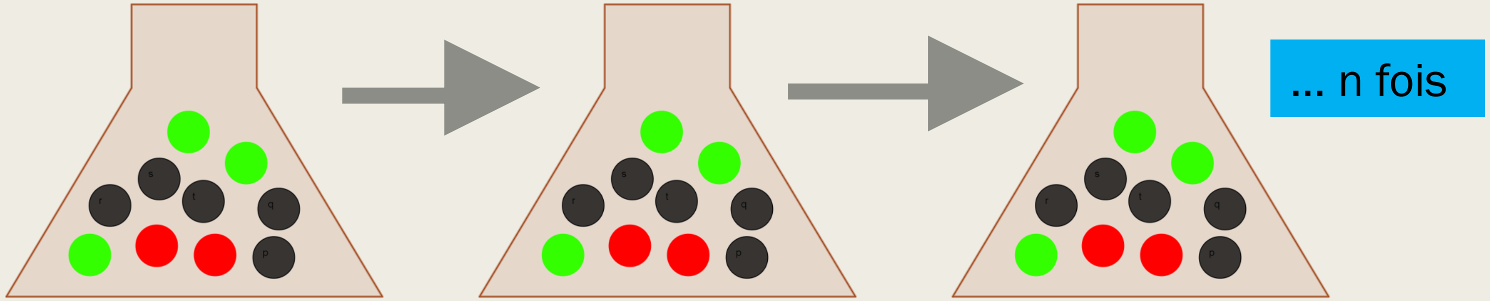
p *: probabilité d'un succès*

q = 1-p : probabilité d'un échec

X : variable aléatoire donnant le nombre de « succès » à l'issue de n essais (de 0 à **n**).

$$P(X = k) = C_n^k p^k (1-p)^{n-k} = C_n^k p^k q^{n-k} \qquad \text{pour} \qquad 0 \leq k \leq n$$

- ***μ=np***
- ***σ² = np(1-p) = npq***



Taux de sondage : $\frac{n}{N}$

(n la taille de l'échantillon, N l'effectif de la population)

Si $\frac{n}{N} \leq \mathbf{0,10}$
on applique la loi **Binomiale.**

Si $\frac{n}{N} \geq \mathbf{0,10}$,
on utilise la loi **Hypergéométrique**

Rappel : $C_n^k = \frac{n!}{k!(n-k)!}$

C- loi HYPERGEOMETRIQUE H(N; D; n)

Soit une population de N individus parmi lesquels D ont un caractère donné. On prélève un échantillon n de cette population N. Les individus de l'échantillon sont tirés simultanément (l'ordre de tirage n'a pas de sens) et sans remise.

- Paramètres: **N** : effectif de la population

  D : nb de personnes présentant le caractère étudié dans la population

  D/N : probabilité p d'avoir le caractère étudié

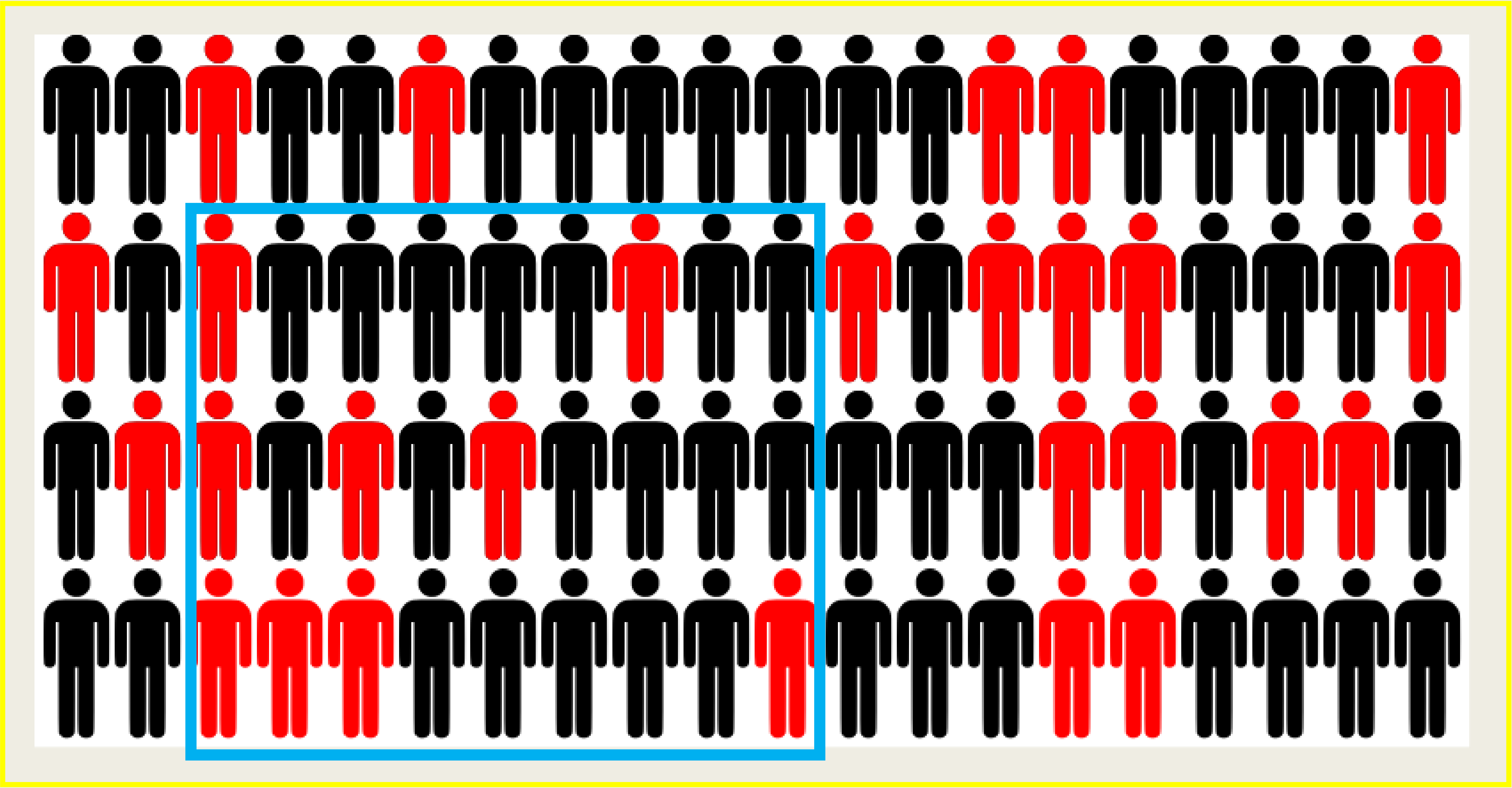
$$\mathbf{P(X = k) = \frac{C_D^k * C_{N-D}^{n-k}}{C_N^n}}$$

- $\mu = \frac{nD}{N} = np$
- $\sigma^2 = npq \times \frac{N-n}{N-1}$

N

D

n



D- loi GEOMETRIQUE G(p)

On répète des épreuves de **Bernoulli jusqu'à l'obtention d'un succès**.

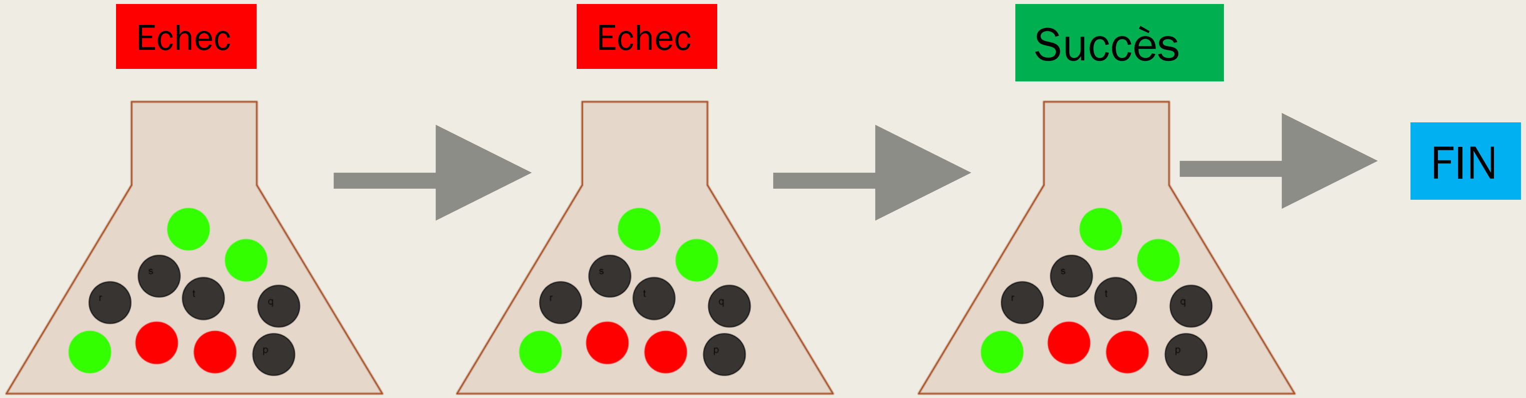
- ❖ Paramètres : **X** : v-a « nb d'essais nécessaires » à l'obtention du 1er succès

  p : probabilité d'un succès

  q : probabilité d'un échec

$$P(X = k) = p(1-p)^{k-1} = pq^{k-1} \qquad \text{pour} \qquad k \in \mathbb{N}^*$$

- ➢ $\mu = \frac{1}{p}$
- ➢ $\sigma^2 = \frac{1-p}{p^2}$



E- loi POISSON P(ʎ)

Utilisée pour déterminer la probabilité qu'un certain nombre d'évènements se réalisent **par unité** de temps (ou de volume, de surface). On la retrouve souvent dans les domaines de la qualité, la sécurité et la fiabilité.

- Paramètres : λ : **taux** moyen avec lequel un évènement particulier se produit en général.

X : variable aléatoire qui donne le nombre d'évènement particulier qui se produisent dans la situation étudiée.

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \qquad \lambda > 0 \qquad k \in \mathbb{N}$$

- **$\mu = \sigma^2 = \lambda$**

QRU SOCRATIVE :

Salle : BIOSTATDUFEU

À propos des lois de probabilité discrète, donnez la réponse vraie:

A- La loi de Bernoulli consiste en n épreuves de loi Binomiale

B- Dans la loi Binomiale, on répète des épreuves de Bernoulli jusqu'à l'obtention du premier succès.

C- La loi Poisson est une expérience aléatoire dont l'issue est un « succès » ou un « échec ».

D- La loi Hypergéométrique est utilisée pour déterminer la probabilité qu'un nombre d'évènements se produisent par unité de temps, de volume ou de surface.

E- Les réponses ABCD sont fausses.

CORRECTION : REPONSE E

À propos des lois de probabilité discrète, donnez la réponse vraie:

A- ~~La loi de Bernoulli consiste en n épreuves de loi Binomiale~~. C'est l'inverse

B- Dans la loi ~~Binomiale~~ Géométrique, on répète des épreuves de Bernoulli jusqu'à l'obtention du premier succès.

C- ~~La loi Poisson~~ Bernoulli est une expérience aléatoire dont l'issue est un « succès » ou un « échec ».

D- La loi ~~Hypergéométrique~~ Poisson est utilisée pour déterminer la probabilité qu'un nombre d'évènements se produisent par unité de temps, de volume ou de surface.

E- Les réponses ABCD sont fausses.

QRU SOCRATIVE :

Salle : BIOSTATDUFEU

À propos des lois de probabilité discrète, donnez la réponse vraie:

A- Si n/N <0,10 on applique la loi Hypergéométrique.

B- Dans la loi de Poisson, **$\mu = \sigma = \lambda$**

C- La loi Binomiale et la loi Hypergéométrique ont toutes les 2 une moyenne **μ=np**

D- La loi Géométrique suit $P(X=k) = p^k(1-p)^{1-k} = p^k q^{1-k} \quad \text{pour} \quad k \in \{0,1\}$

E- Les réponses ABCD sont fausses.

CORRECTION : REPONSE C

À propos des lois de probabilité discrète, donnez la réponse vraie:

A- Si n/N <0,10 on applique la loi ~~Hypergéométrique~~ Binomiale.

B- Dans la loi de Poisson, **$\mu = \sigma^2 = \lambda$**

C- La loi Binomiale et la loi Hypergéométrique ont toutes les 2 une moyenne **$\mu = np$**

D- La loi Géométrique suit $P(X = k) = p(1-p)^{k-1} = pq^{k-1}$ pour $k \in \mathbb{N}^*$

E- Les réponses ABCD sont fausses.

QRU SOCRATIVE :

Salle : BIOSTATDUFEU

On compte en moyenne 12 transplantations toutes les 2semaines à Nice, quelle est la probabilité de n'en compter que 3 en 1 semaine ?

A- On utilise une loi de Poisson : $P(\mathrm{X} = 3) = 36\mathrm{e}^{-6}$

B- On utilise une loi de Poisson : $P(\mathrm{X} = 3) = \frac{12^3 e^{-12}}{3!}$

C- On utilise une loi Hypergéométrique : $P(X = 3) = \frac{C_{12}^3 C_{12}^9}{C_9^3}$

D- On utilise une loi de Hypergéométrique : $P(X = 3) = 2$

E- Les réponses ABCD sont fausses.

CORRECTION : REPONSE A

On compte en moyenne 12 transplantations toutes les 2 semaines à Nice, quelle est la probabilité de n’en compter que 3 en 1 semaine ?

$$P(X=3)=\frac{\lambda^k e^{-\lambda}}{k!}=\frac{6^3\mathrm{e}^{-6}}{3!}=\frac{6\times 36\mathrm{e}^{-6}}{3\times 2}=\mathbf{36\ e^{-6}}$$

A- On utilise une loi de Poisson :$P(\mathrm{X}=3)=36\mathrm{e}^{-6}$

B- On utilise une loi Hypergéométrique :$P(\mathrm{X}=3)=\frac{12^3 e^{-12}}{3!}$

C- On utilise une loi Hypergéométrique :$P(X=3)=\frac{C_{12}^3 C_{12}^9}{C_9^3}$

D- On utilise une loi de Poisson : $P(X=3)=2$

E- Les réponses ABCD sont fausses.

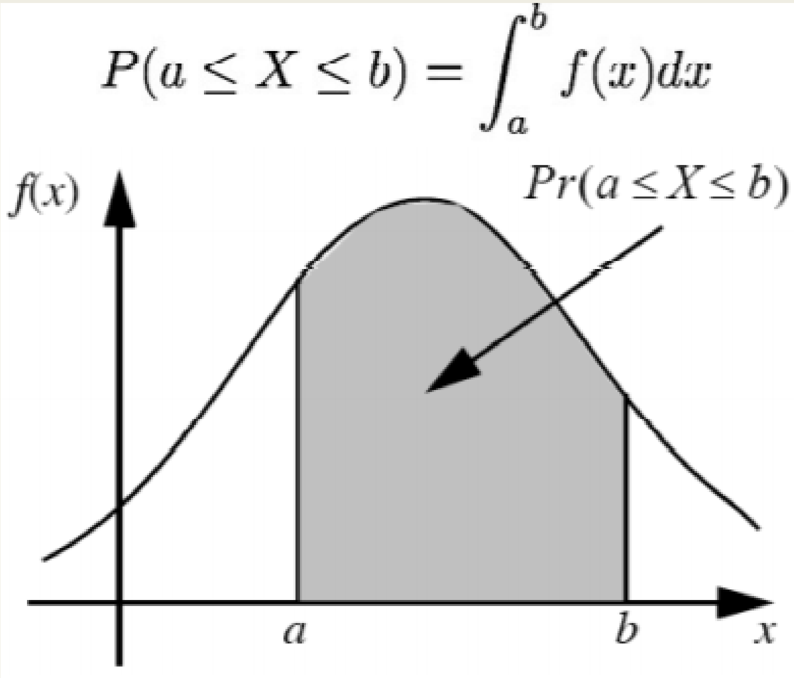
IV- VARIABLE ALEATOIRE CONTINUE

Ce qui caractérise une variable aléatoire continue, c'est qu'elle a une probabilité nulle d'être égale à un nombre donné

$P(X=k) = 0$

On utilisera donc des intervalles : $P(a \leq X \leq b) \neq 0$

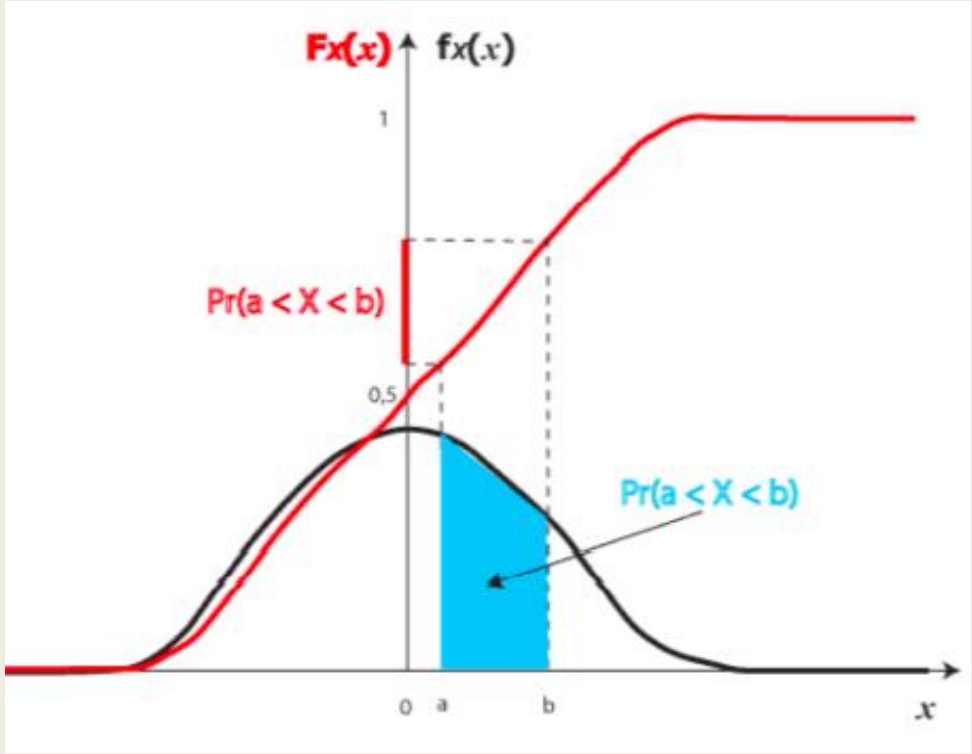
- **La densité de probabilité :** C'est une fonction utilisée pour définir la loi de probabilité de X. $P(a \leq X \leq b)$ est l'aire sous la courbe.



•La Fonction de Répartition :

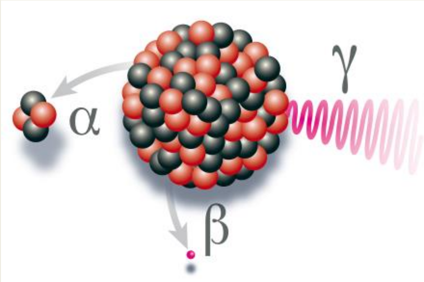
Elle est toujours **croissante**, **monotone** et **continue**.

Partant de 0 pour x → -∞ . Atteignant 1 pour x → +∞ . (en rouge la fonction de répartition, en noir la fonction de densité)



V- LOIS DE PROBABILITE CONTINUE

A- loi EXPONENTIELLE E(λ)



On l'utilise dans des situations où le « risque instantané » de décès ou « taux de défaillance » est constant. (durée de vie d'un composant…)

- **Fonction de densité :** avec $\lambda > 0$ et $x \geq 0$ $$\mathbf{f}(x) = \lambda \mathbf{e}^{-\lambda x}$$

  - *Paramètres :* λ = *taux de défaillance instantané*

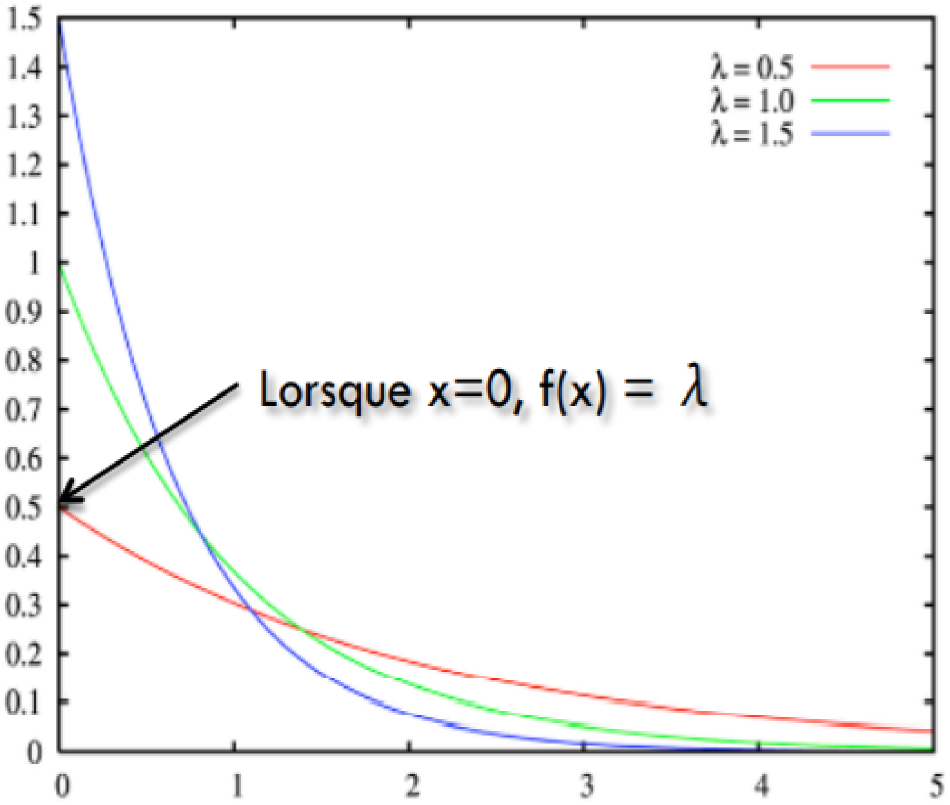
    $\mu = 1 / \lambda$

    $\sigma^2 = 1 / \lambda^2$

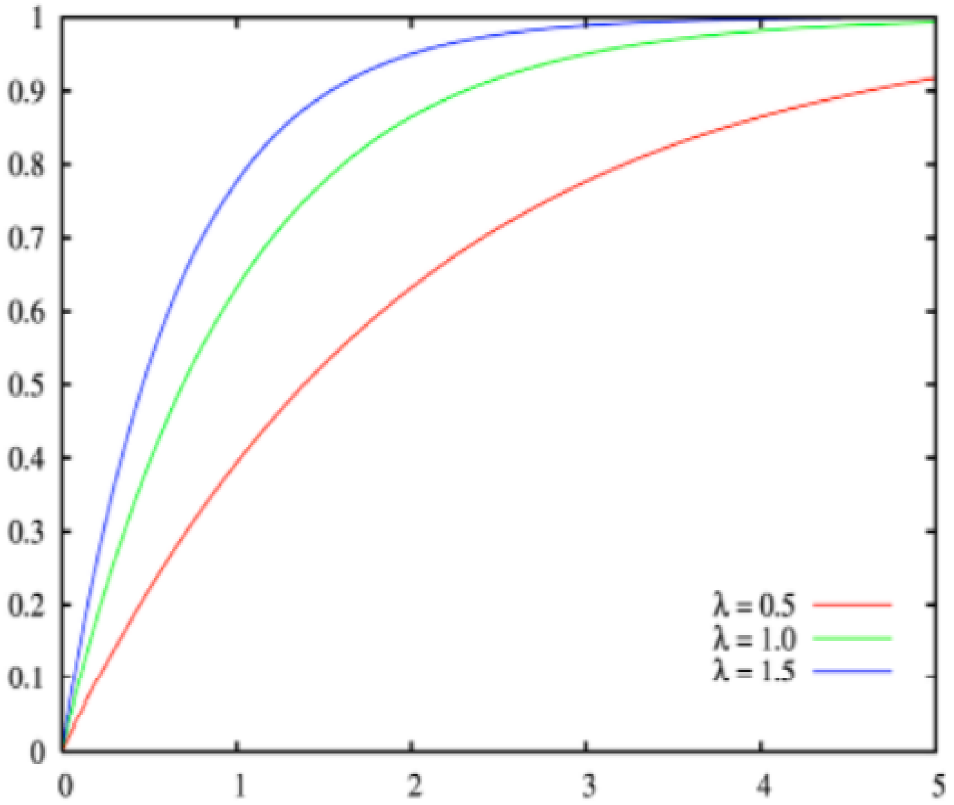
- **Fonction de répartition:** $$f(x) = P(X \leq x) = \int_0^x \lambda \mathbf{e}^{-\lambda t}\, \mathbf{d}t = \mathbf{1} - \mathbf{e}^{-\lambda x}$$

Si un évènement se réalise selon une loi de **Poisson** de paramètre λ, le **temps entre deux réalisations consécutives** de l'évènement considéré est distribué selon une **loi exponentielle de paramètre 1/λ.**

Densité de probabilités (f):



Fonction de répartition (F):



B- loi <u>UNIFORME</u> U(a;b)

- **<u>Fonction de densité</u>** : $f(x) = \frac{\mathbf{1}}{\boldsymbol{b-a}}$ si $x \in [a, b]$

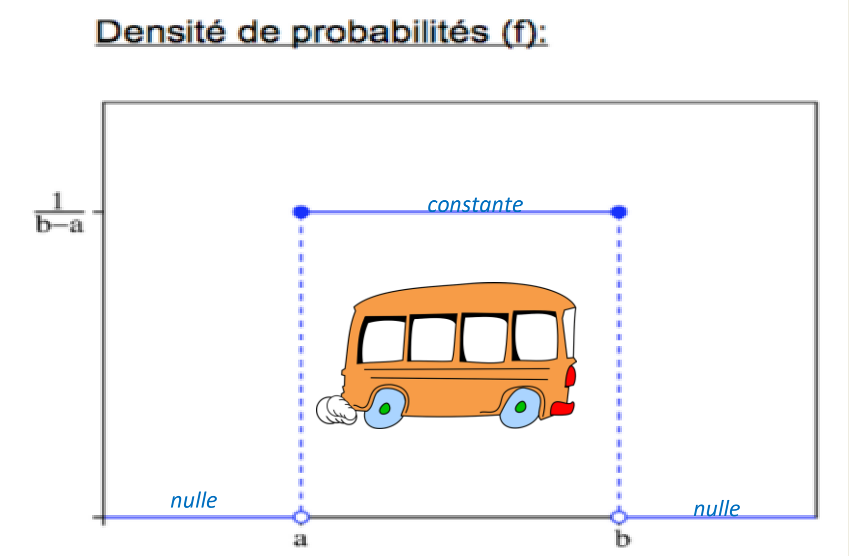
  $f(x) = 0 \qquad$ si $x \notin [a, b]$

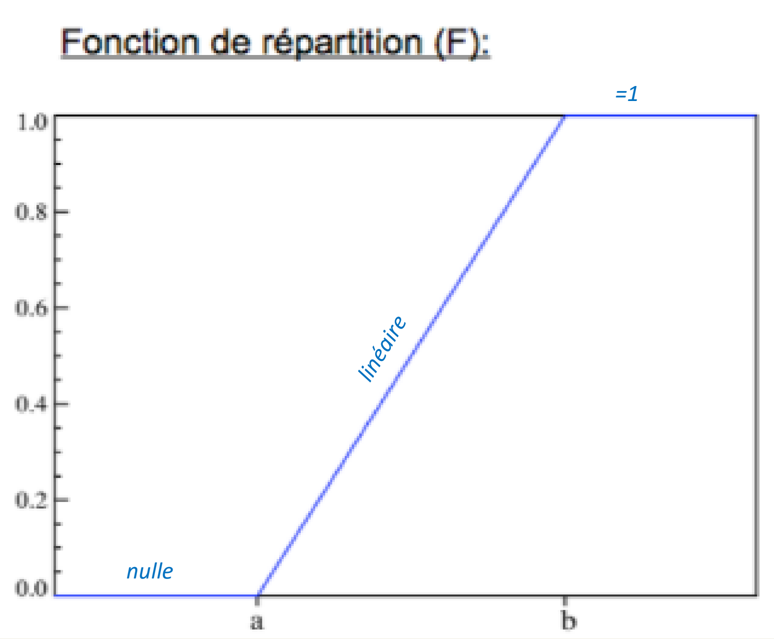
  avec $\lambda > 0$ et $x \geq 0$

❖ Paramètres : intervalle [a,b] ∈ $\mathbb{R}$

- $\boldsymbol{\mu} = \frac{\boldsymbol{a+b}}{\mathbf{2}}$
- $\boldsymbol{\sigma^2} = \frac{(\boldsymbol{b-a})^2}{\mathbf{12}}$

Toutes les valeurs ont la même probabilité.



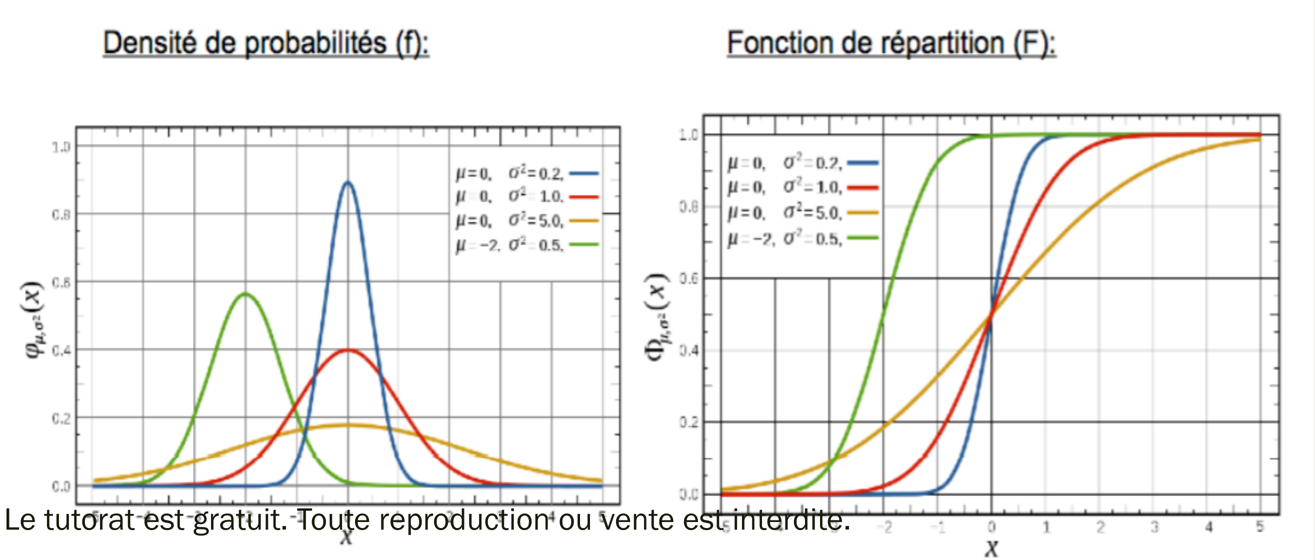


C- loi NORMALE U(μ; σ)

- Fonction de densité : $f(x) = \frac{1}{\sigma\sqrt{2\Pi}} e^{-\frac{(x-\mu)^2}{\sigma^2}}$ pour $-\infty \geq x \geq +\infty$

❖ Paramètres : **μ** et **σ** , respectivement moyenne et écart-type de X.

La **densité** de probabilité d'une v-a normale est **symétrique** autour de μ et a deux points d'inflexion aux abscisses μ - σ et μ + σ. Son utilisation est très vaste.



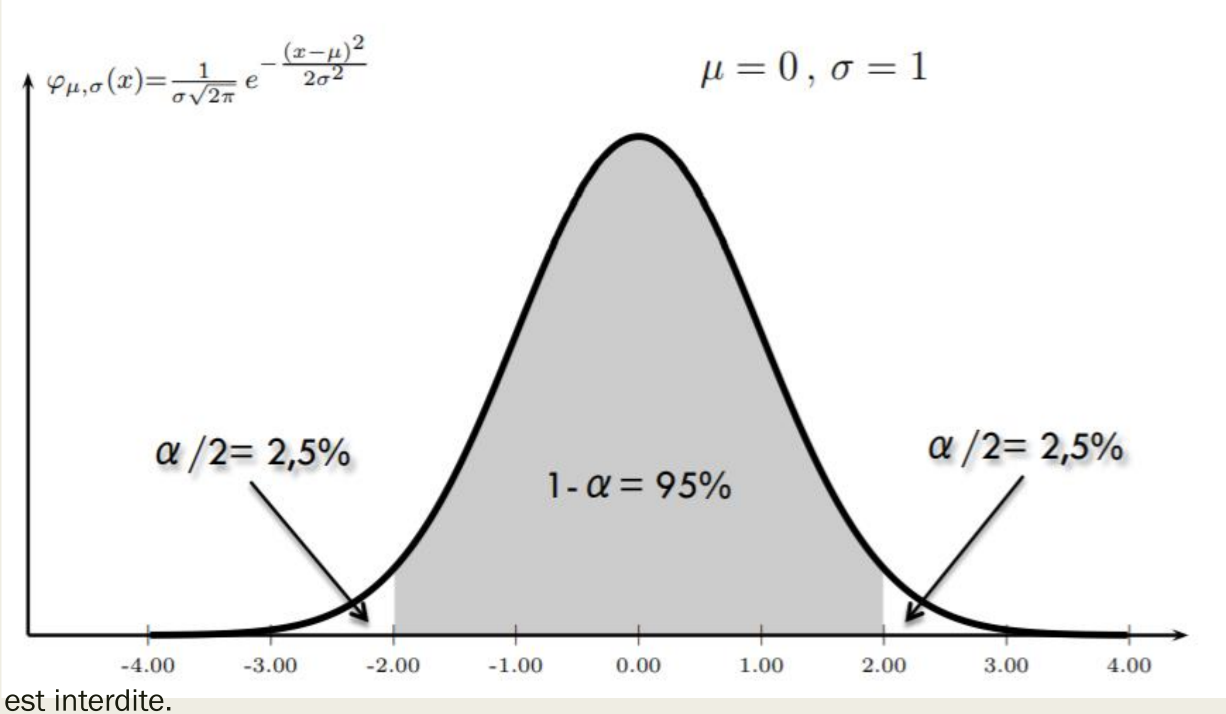
D- loi NORMALE CENTREE REDUITE N(0;1)

La loi normale centrée réduite est une loi normale de moyenne 0 et de variance 1.

$$z = \frac{X-\mu}{\sigma}$$

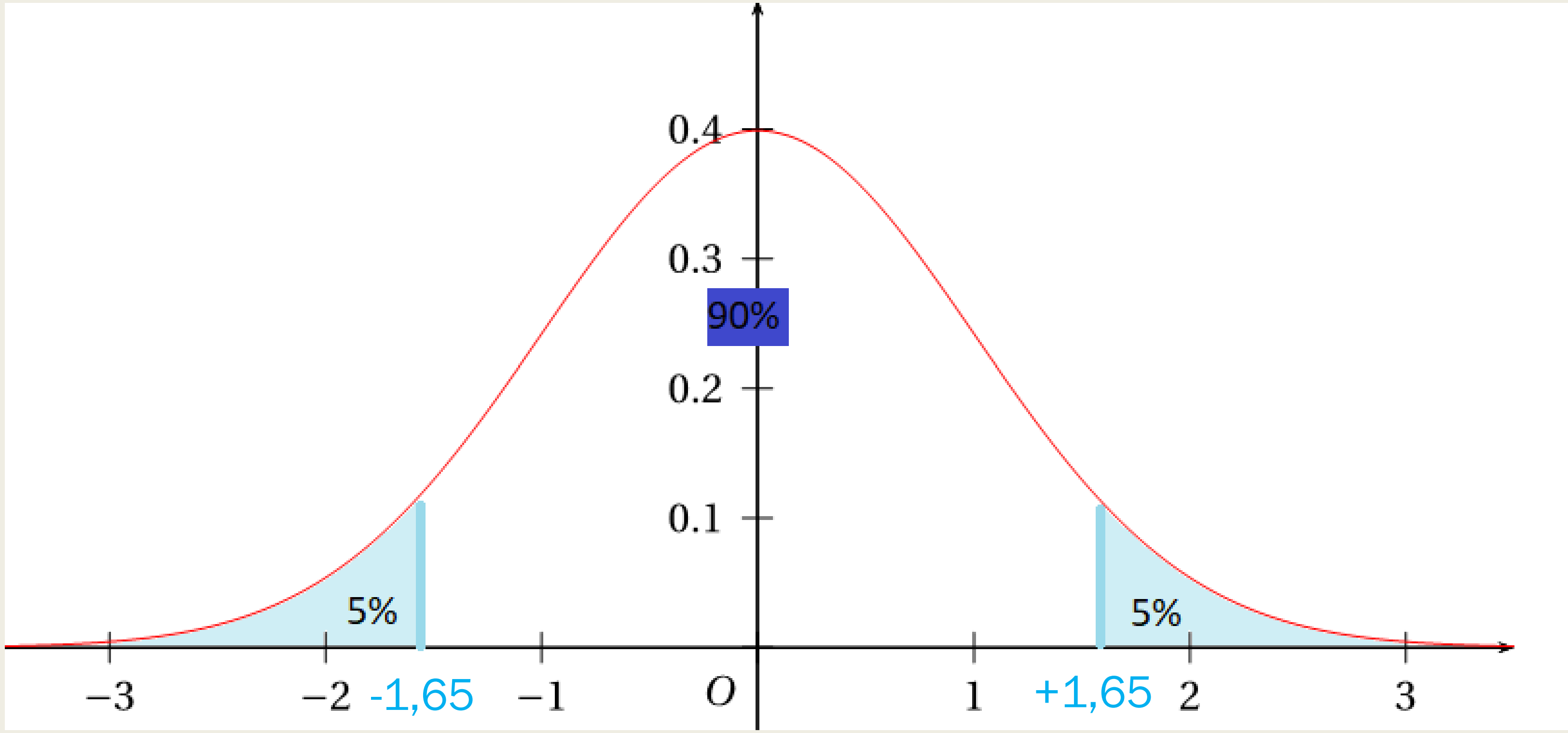
Centrée → Autour de la moyenne μ = 0

- Réduite → Ayant une variance σ = 1
- N(μ ; σ) => N(0 ; 1)



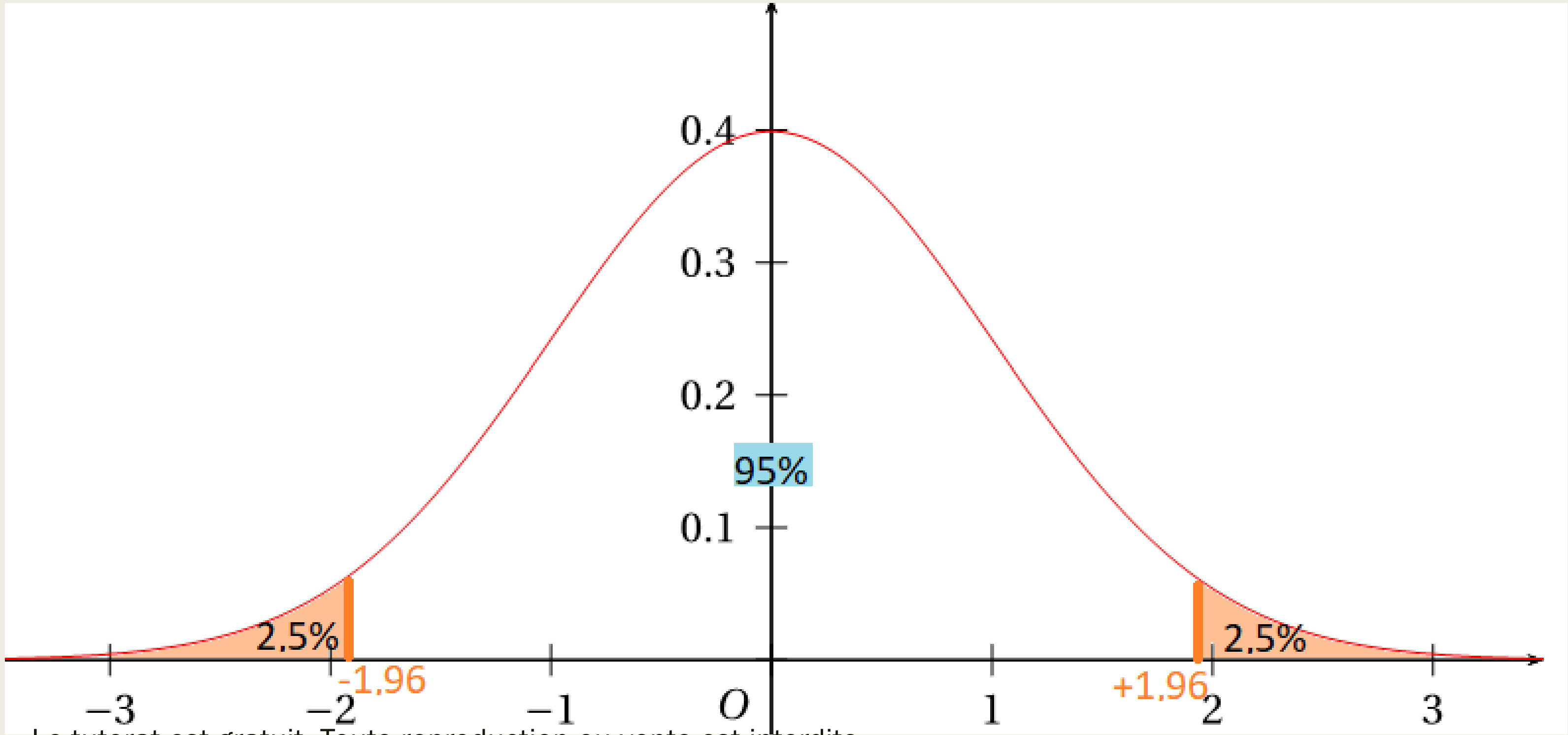
Valeurs limites importantes à savoir

- il y a 10 chances sur 100 pour que $X < \mu - 1,65\sigma$ ou $X > \mu + 1,65\sigma$
- il y a 5 chances sur 100 pour que $X < \mu - 1,96\sigma$ ou $X > \mu + 1,96\sigma$
- il y a 1 chance sur 100 pour que $X < \mu - 2,58\sigma$ ou $X > \mu + 2,58\sigma$
- il y a 1 chance sur 1000 pour que $X < \mu - 3,30\sigma$ ou $X > \mu + 3,30\sigma$



Valeurs limites importantes à savoir

- il y a 10 chances sur 100 pour que $X < \mu - 1,65\sigma$ ou $X > \mu + 1,65\sigma$
- il y a 5 chances sur 100 pour que $X < \mu - 1,96\sigma$ ou $X > \mu + 1,96\sigma$
- il y a 1 chance sur 100 pour que $X < \mu - 2,58\sigma$ ou $X > \mu + 2,58\sigma$
- il y a 1 chance sur 1000 pour que $X < \mu - 3,30\sigma$ ou $X > \mu + 3,30\sigma$



APPROXIMATIONS :

LOIS	CONDITIONS	CONSEQUENCE
la loi **BINOMIALE** peut être approximée par une loi POISSON	Si **N>50** **p≤0,10** **np ≤ 5**	B(n;p) → P(λ=np)
La loi **BINOMIALE** peut être approximée par une loi NORMALE	Si **np ≥ 5** **nq ≥ 5**	B(n;p) → N(np;$\sqrt{npq}$)
la loi **POISSON** peut être approximée par une loi NORMALE	Si **λ > 25**	P(λ) → N(λ;$\sqrt{\lambda}$)

QRU SOCRATIVE :

Salle : BIOSTATDUFEU

À propos des lois de probabilité continues, donnez la réponse vraie:

A- Ce qui caractérise une v.a continue c’est qu’elle a une probabilité nulle d’être égale à un nombre donné : $P(X=k)=0$

B- La fonction de répartition est toujours croissante, monotone et discontinue.

C- La représentation de la densité de probabilité de la loi normale est asymétrique.

D- La loi Normale est utilisée dans des situations où le risque instantané de décès est constant.

E- Les propositions A,B,C et D sont fausses.

CORRECTION : REPONSE A

À propos des lois de probabilité continues, donnez la réponse vraie:

A- Ce qui caractérise une v.a continue c'est qu'elle a une probabilité nulle d'être égale à un nombre donné : $P(X=k)=0$

B- La fonction de répartition est toujours croissante, monotone et ~~discontinue~~.

C- La représentation de la densité de probabilité de la loi normale est ~~asymétrique~~.

D- La loi ~~Normale~~ Exponentielle est utilisée dans des situations où le risque instantané de décès est constant.

E- les réponses ABCD sont fausses.

QRU SOCRATIVE : PACES

Salle : BIOSTATDUFEU

Dans la promo de Paces, on compte 400 étudiants. Parmi eux, 30% sont des hommes. On tire indépendamment et aléatoirement 20 P1. Donnez la réponse vraie.

A- On est dans le domaine d'application de la loi Hypergéométrique.

B- La probabilité de tirer 5 hommes sur les 20 est de $P(X = 5) = C_{20}^{5} \times 0{,}3^{5} \times 0{,}7^{15}$

C- La probabilité de tirer 5 femmes est de $P(X = 5) = C_{20}^{5} \times 0{,}3^{5} \times 0{,}7^{15}$

D- On ne peut pas répondre, il manque une donnée.

E- les réponses ABCD sont fausses.

CORRECTION : REPONSE B

Dans la promo de Paces, on compte 400 étudiants. Parmi eux, 30% sont des hommes. On tire indépendamment et aléatoirement 20 P1. Donnez la réponse vraie.

A- On est dans le domaine d'application de la loi ~~Hypergéométrique~~.

B- La probabilité de tirer 5 hommes sur les 20 est de $P(X = 5) = C_{20}^{5} \times 0{,}3^{5} \times 0{,}7^{15}$

C- La probabilité de tirer 5 femmes est de $P(X = 5) = C_{20}^{5} \times 0{,}3^{5} \times 0{,}7^{15}$

D- On ne peut pas répondre, il manque une donnée.

E- les réponses ABCD sont fausses.

$$P(X = k) = C_n^k \times p^k q^{n-k}$$

Taux de sondage = 5% < 0,10 Donc on utilise la Binomiale

A VOS QUESTIONS !