

LE RESUME de « probabilités et statistiques »

Rédigé par :

Kevin BARZA

1^{ere} CE – Semestre 2

2015-2016

MES

Sommaire réduit

Chapitre 1 : Dénombrement	1
Chapitre 2 : Calculs des probabilités	2
1. Introduction	3
2. Espace probabilisable	3
3. Espace probabilisé	3
4. Calcul de probabilité.....	4
5. Probabilité conditionnelle	4
6. Evènements indépendants	4
Chapitre 3 : Variables aléatoires, distribution de probabilité	5
1. Variables aléatoires	5
2. Distributions de probabilités et caractéristiques	5
Chapitre 4 : Lois de distribution usuelles	8
1. Distributions discrètes.....	9
2. Distributions continues.....	12
Chapitre 5 : Transformations de variables	16
1. Moments et fonctions caractéristiques de variables aléatoires	16
2. Transformations de variables	17
Chapitre 6 : Modes de convergence des variables aléatoires (hors programme)	
Chapitre 7 : Estimation	18
1. Vocabulaire statistique	18
2. Exemples de statistiques	19
3. Distributions d'échantillonnage	19
4. Estimations	21
Chapitre 8 : Tests d'Hypothèses	24
1. Vocabulaire	24
2. Démarche à suivre lors d'un test d'hypothèse sur des paramètres	25
3. Tests usuels sur des paramètres	25
4. Tests sur d'autres caractéristiques	26
Chapitre 9 : Régression linéaire (pas encore disponible)	

Sommaire détaillé

Chapitre 1 : Dénombrement	1
Chapitre 2 : Calculs des probabilités	2
1. Introduction	
1.1 Expérience	
1.2 Univers des possibles	
1.3 Evènement	
1.4 Terminologie	
1.5 Système complet d'évènements ou partition de Ω	
2. Espace probabilisable	
2.1 Tribu	
2.2 Espace probabilisable ou mesurable	
3. Espace probabilisé	
3.1 Mesure	
3.2 Espace mesuré	
3.3 Probabilité	
3.4 Espace probabilisé	
4. Calcul de probabilité	
5. Probabilité conditionnelle	
6. Evènements indépendants	
Chapitre 3 : Variables aléatoires, distribution de probabilité	5
1. Variables aléatoires	
1.1. Variable aléatoire	
1.2. Vecteur aléatoire ou système de variables aléatoire	
2. Distributions de probabilités et caractéristiques	
2.1. Distribution de probabilité d'une seule variable aléatoire X	
2.2. Fonction de répartition	
2.3. Distribution jointe d'un système de deux variables aléatoires X et Y	
2.4. Distribution marginale	
2.5. Distribution conditionnelle	
2.6. Indépendance	
2.7. Indépendance mutuelle	
2.8. Esperance d'une variable aléatoire X	
2.9. Esperance suite à un changement d'une variable aléatoire à une nouvelle	
2.10. Esperance suite à un changement de deux variables aléatoires en une nouvelle	
2.11. Propriétés de l'esperance	
2.12. Variance et écart-type	
2.13. Propriété de la variance	
2.14. Covariance	
2.15. Propriétés de la covariance	

- 2.16. Relation entre variance et covariance
- 2.17. Propriété de la relation
- 2.18. Coefficient de corrélation linéaire r

Chapitre 4 : Lois de distribution usuelles8

- 1. Distributions discrètes
 - 1.1. Uniforme
 - 1.2. Bernoulli
 - 1.3. Binomiale
 - 1.4. Multinomiale
 - 1.5. Hypergéométrique
 - 1.6. Hypergéométrique multi-variée
 - 1.7. Binomiale négative
 - 1.8. Géométrique
 - 1.9. Poisson
- 2. Distributions Continues
 - 2.1. Uniforme
 - 2.2. Normale (en général)
 - 2.3. Normale standard
 - 2.4. Gamma
 - 2.5. Exponentielle
 - 2.6. Béta
 - 2.7. Log-normale
 - 2.8. Chi 2
 - 2.9. Student

Chapitre 5 : Transformations de variables16

- 1. Moments et fonctions caractéristiques de variables aléatoires
 - 1.1. Moment d'une variable aléatoire
 - 1.2. Fonction caractéristique
 - 1.3. Relation entre moments et fonctions caractéristiques
- 2. Transformations de variables
 - 2.1. Cas d'une seule variable aléatoire
 - 2.2. Cas de deux variables aléatoires

Chapitre 6 : Modes de convergence des variables aléatoires (hors programme)

Chapitre 7 : Estimation18

- 1. Vocabulaire statistique
 - 1.1. Observation
 - 1.2. Population
 - 1.3. Echantillon
 - 1.4. Taille d'une population (ou d'une échantillon)
 - 1.5. Echantillonnage
 - 1.6. Echantillonnage aléatoire simple
 - 1.7. Statistique

2. Exemples de statistiques
 - 2.1. Moyenne
 - 2.2. Médiane
 - 2.3. Mode
 - 2.4. Etendue
 - 2.5. Variance empirique
 - 2.6. Variance empirique modifiée
 - 2.7. Chi-2
 - 2.8. Proportion
3. Distributions d'échantillonnage
 - 3.1. Définition
 - 3.2. Distribution de Chi-2
 - 3.3. Distribution de la moyenne
 - 3.4. Distribution de Student
 - 3.5. Distribution F
 - 3.6. Récapitulatif
4. Estimations
 - 4.1. Estimation ponctuelle
 - 4.2. Estimation par intervalle de confiance
 - 4.3. Estimation par maximum de vraisemblance
 - 4.4. Estimation par la méthode des moments

Chapitre 8 : Tests d'Hypothèses24

1. Vocabulaire
 - 4.1. Hypothèse statistique
 - 4.2. Hypothèse nulle H_0 et hypothèse alternative H_1
 - 4.3. Test d'hypothèse (ou test statistique)
 - 4.4. Erreurs
 - 4.5. Niveau de signifiante (ou seuil de signification) α
 - 4.6. Valeur critique
 - 4.7. P-valeur
 - 4.8. Statistique de test (notée ST ou T)
2. Démarche à suivre lors d'un test d'hypothèse sur des paramètres
3. Tests usuels sur des paramètres
 - 3.1. Moyenne
 - 3.2. Différence entre deux moyennes
 - 3.3. Proportion
 - 3.4. Variance
4. Tests sur d'autres caractéristiques

Chapitre 9 : Régression linéaire (pas encore disponible)

Chapitre 1 : Dénombrement

- Un arrangement : une représentation d'un ensemble d'éléments en tenant compte de l'ordre dans lequel ils ont été placés (ou énumérés).
Exemple : les ensembles {1 ;2} et {2 ;1} sont différents, comptés deux fois.
- Une combinaison : une représentation d'un ensemble d'éléments sans tenir compte de l'ordre dans lequel ils ont été placés (ou énumérés).
Exemple : les ensembles {1 ;2} et {2 ;1} sont identiques, comptés une seule fois.
- Une répétition : la présence d'un même élément plus qu'une fois dans un sous-ensemble.
Exemple : soit une urne contenant 3 boules : blanche, noire et rouge. Supposons qu'on tire une boule et qu'elle soit rouge. Si on la remet et on tire une nouvelle fois, il se peut qu'on ait une rouge une autre fois, d'où la répétition. Si on tire une nouvelle sans pour autant remettre la première, il n'y a donc plus de répétition.
On associe généralement la répétition à un tirage avec remise, et la non répétition à la non remise.

	Répétition	Ordre	Formule	Remarques
1. P-liste (ou arrangement avec répétition)	OUI	OUI	$\forall n, p \in \mathbb{N}$, On a n^p façons d'arranger p éléments à partir de n éléments.	
2. Arrangement (sans répétition)	NON	OUI	$\forall n, p \in \mathbb{N} \text{ tq } n \geq p$, On a $A_n^p = \frac{n!}{(n-p)!}$ façons d'arranger p éléments à partir de n éléments.	Si $n = p$, on parle de permutation : $P_n = n!$ Si parmi les n , on a n_1 éléments semblables, n_2 autres semblables, ..., n_k autres semblables tq $\sum n_k = n$, on parle alors de $\frac{n!}{n_1!n_2!\dots n_k!}$ permutations distincts.
3. Combinaison (sans répétition)	NON	NON	$\forall n, p \in \mathbb{N} \text{ tq } n \geq p$, On a $C_n^p = \binom{n}{p} = \frac{n!}{p!(n-p)!}$ façons de combiner p éléments à partir de n éléments.	Si parmi n éléments, on a n_1 éléments semblables, n_2 autres semblables, ..., n_k autres semblables, tq $\sum n_k = n$, on parle alors de $\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1!n_2!\dots n_k!}$ combinaisons distincts.
4. Combinaison (avec répétition)	OUI	NON	$\forall n, p \in \mathbb{N}$, on a : $C_{n+p-1}^p = \binom{n+p-1}{p}$ façons de combiner p éléments à partir de n éléments.	

Chapitre 2 : Calcul des probabilités

1. Introduction

1.1. Expérience

Dans la vie pratique, il existe 2 types d'expérience :

- Déterministe : dont on peut connaître le résultat avant sa réalisation en utilisant toutes les informations qui sont à notre disposition.
- Aléatoire : dont le résultat ne dépend que du hasard et donc imprévisible. C'est à ce type d'expérience qu'on s'intéresse en probabilité.

1.2. Univers des possibles

Appelé aussi tout simplement univers et noté Ω , il représente l'ensemble de tous les résultats possible d'une expérience.

1.3. Evènement

Un évènement est un sous ensemble de l'univers Ω .

1.4. Terminologie

- Evènement élémentaire : un singleton $\{w\}$
- Evènement impossible : Φ
- Evènement certain : Ω
- Evènement contraire de A : \bar{A}
- Evènement « A et B » : $A \cap B$
- Evènement « A ou B » : $A \cup B$
- Evènements A et B incompatibles : $A \cap B = \Phi$

1.5. Système complet d'évènements ou partition de Ω

C'est une famille d'évènements $\{A_i\}$ tel que :

- $\forall i \neq j, A_i \cap A_j = \Phi$ c'est à dire que les évènements sont 2 à 2 incompatibles.
- $\bigcup A_i = \Omega$.

2. Espace probabilisable

2.1. Tribu ou σ -algèbre

Définition :

Une tribu sur Ω est un sous-ensemble T de l'ensemble des parties de Ω ($P(\Omega)$) telle que :

- $\Omega \in T$
- $A \in T \rightarrow \bar{A} \in T$
- Pour toute famille dénombrable (A_i) d'éléments de T , on a $\cup A_i \in T$

Propriétés :

- $\emptyset \in T$
- $A, B \in T \rightarrow A \cup B \in T$
- Pour toute famille dénombrable (A_i) d'éléments de T , on a $\cap A_i \in T$
- $A, B \in T \rightarrow A \cap B \in T$

2.2. Espace probabilisable ou mesurable

On appelle espace probabilisable ou mesurable tout ensemble Ω muni d'une tribu T .

On le note par le couple (Ω, T) .

3. Espace probabilisé

3.1. Mesure

Soit (Ω, T) un espace mesurable. On appelle mesure sur (Ω, T) une application $f: T \rightarrow R_+$ tq:

- Pour toute famille dénombrable (A_i) d'éléments de T 2 à 2 disjoints, on a: $f(\cup A_i) = \sum f(A_i)$.

3.2. Espace mesuré

On appelle espace mesuré un espace mesurable (Ω, T) munie d'une mesure f .

On le note par le couple (Ω, T, f) .

3.3. Probabilité

Une probabilité n'est d'autre qu'une mesure dont le domaine de définition peut être restreint sur $[0; 1]$ c'est à dire $p: T \rightarrow R_+$, qui en plus de la condition précédente, doit vérifier 2 autres :

- $\forall A \in T, 0 \leq p(A) \leq 1$
- $p(\Omega) = 1$

3.4. Espace probabilisé

On appelle espace probabilisé un espace probabilisable (Ω, T) munie d'une probabilité p .

On le note par le couple (Ω, T, p) .

4. Calcul de probabilité

Théorème :

Soient A et B deux évènements quelconques , alors : $p(A \cup B) = p(A) + p(B) - p(A \cap B)$

Remarque : Evidemment si A et B sont incompatibles alors $p(A \cup B) = p(A) + p(B)$

Théorème généralisé :

Soient (A_n) une famille d'évènements quelconques, alors :

$$p\left(\bigcup_{k=1}^n A_k\right) = \sum_{k=1}^n p(A_k) - \sum_{i < j} p(A_i \cap A_j) + \sum_{i < j < k} p(A_i \cap A_j \cap A_k) + \dots + (-1)^{n+1} p(A_1 \cap A_2 \dots \cap A_n)$$

Remarque : Evidemment si les A_i sont incompatibles alors $p(\bigcup_{k=1}^n A_k) = \sum_{k=1}^n p(A_k)$

En particulier : Pour les évènements A et \bar{A} , on a : $1 = p(A) + p(\bar{A}) \rightarrow p(A) = 1 - p(\bar{A})$

Théorème :

Soit A et D deux évènements. On a : $p(A) = p(A \cap D) + p(A \cap \bar{D})$

5. Probabilité conditionnelle

Définition :

Soit un évènement A tel que $p(A) \neq 0$. On définit la probabilité de B sachant A par :

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

Remarque : il en découle que $p(A \cap B) = p(A) \times p(B|A) = p(B) \times p(A|B)$

Théorème des probabilités totales :

Soient un évènement A et un système complet d'évènements (B_i) . On a alors :

$$P(A) = \sum_i p(A \cap B_i) = \sum_i p(B_i) \times p(A|B_i)$$

Théorème de Bayes:

Soient un évènement A et un système complet d'évènements (B_i) tel que $p(B_i) \neq 0$. On a :

$$P(B_j|A) = \frac{p(B_j \cap A)}{\sum p(A \cap B_i)} = \frac{p(B_j) \times p(A|B_j)}{\sum_i p(B_i) \times p(A|B_i)}$$

6. Evènements indépendants

• On dit que deux évènements A et B sont indépendants si $p(A \cap B) = p(A) \times p(B)$.
Il en découle que si A et B sont indépendants alors $p(A|B) = p(A)$ et $p(B|A) = p(B)$.

• On dit que n évènements A_1, A_2, \dots, A_n sont mutuellement indépendants si :

$$p\left(\bigcap A_i\right) = p(A_1) \times p(A_2) \dots \times p(A_n)$$

Chapitre 3 : Variables aléatoires et distributions de probabilités

1. Variables aléatoires

Soit Ω l'univers des possibles

1.1. Variable aléatoire

Une variable aléatoire X est une application $X : \Omega \rightarrow \mathbb{R}$ c'est à dire qui associe à chaque élément de Ω un certain nombre réel.

Elle est dite :

- Discrète si elle ne prend que des valeurs discrètes (des nombres), dont l'ensemble de ces valeurs est au plus dénombrable (c'est à dire $X(\Omega)$ en bijection avec \mathbb{N}).
- Continue si elle prend des valeur sous forme d'intervalle de \mathbb{R} .

1.2. Vecteur aléatoire ou système de variables aléatoires

Un vecteur aléatoire X de dimension n est une application $X : \Omega \rightarrow \mathbb{R}^n$ tel que

$X = (X_1, X_2, \dots, X_n)$ à pour coordonnées n variables aléatoires qui sont les X_i .

On parle en d'autres termes d'un système de n variables aléatoires X_1, X_2, \dots, X_n .

2. Distributions de probabilités et caractéristiques

Cas :	Discret	Continue
1. Distribution de probabilité d'une seule variable aléatoire X	Appelé aussi loi de probabilité, c'est une application : $f: \mathbb{R} \rightarrow [0; 1]$ $x \rightarrow f(x)$ Vérifiant : <ul style="list-style-type: none"> • $\sum_x f(x) = 1$ • $p(X = x) = f(x) \leq 1$ 	Appelé aussi densité de probabilité, c'est une application : $f: \mathbb{R} \rightarrow \mathbb{R}_+$ $x \rightarrow f(x)$ Vérifiant : <ul style="list-style-type: none"> • $\int_{-\infty}^{+\infty} f(x)dx = 1$ • $p(a < X < b) = \int_a^b f(x)dx \leq 1$
2. Fonction de répartition	C'est une application : $F: \mathbb{R} \rightarrow [0; 1]$ $x \rightarrow P(X \leq x) = \sum_{k \leq x} f(k)$	C'est une application : $F: \mathbb{R} \rightarrow [0; 1]$ $x \rightarrow P(X \leq x) = \int_{-\infty}^x f(x)dx$
Remarque :	$p(a < X < b) = F(b) - F(a)$	$p(a < X < b) = F(b) - F(a)$ Et $f(x) = \frac{dF(x)}{d(x)}$ si la dérivée existe

3. Distribution jointe d'un système de deux variables aléatoires X et Y	C'est une application : $f: \mathbf{R}^2 \rightarrow [0; 1]$ $(x, y) \rightarrow f(x, y)$ Vérifiant : <ul style="list-style-type: none"> $\sum_x \sum_y f(x, y) = 1$ $p(X = x; Y = y) = f(x, y) \leq 1$ 	C'est une application : $f: \mathbf{R}^2 \rightarrow \mathbf{R}_+$ $(x, y) \rightarrow f(x, y)$ Vérifiant : <ul style="list-style-type: none"> $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x, y) dx dy = 1$ $p((X, Y) \in S) = \iint_S f(x, y) dx dy \leq 1$ S étant une surface du plan xy .
4. Distribution marginale	Celle de X est défini par : $g: \mathbf{R} \rightarrow [0; 1]$ $x \rightarrow g(x) = \sum_y f(x, y)$ Celle de Y est défini par : $h: \mathbf{R} \rightarrow [0; 1]$ $y \rightarrow h(y) = \sum_x f(x, y)$	Celle de X est défini par : $g: \mathbf{R} \rightarrow [0; 1]$ $x \rightarrow g(x) = \int_{-\infty}^{+\infty} f(x, y) dy$ Celle de Y est défini par : $h: \mathbf{R} \rightarrow [0; 1]$ $y \rightarrow h(y) = \int_{-\infty}^{+\infty} f(x, y) dx$
5. Distribution conditionnelle	Celle de Y sachant que $X = x$ est défini par : $f(y x) = \frac{f(x, y)}{g(x)} \quad \text{avec } g(x) \neq 0$ Celle de X sachant que $Y = y$ est défini par : $f(x y) = \frac{f(x, y)}{h(y)} \quad \text{avec } h(y) \neq 0$	
6. Indépendance	Soient $X, Y, f(x, y), g(x), h(y)$ comme avant. On dit que X et Y sont indépendantes ssi : $\forall x, y \in \mathbf{R}, \quad f(x, y) = g(x) \times h(y)$	
7. Indépendance mutuelle	Soient X_1, X_2, \dots, X_n n variables aléatoires de distribution de probabilité $f(x_1, x_2, \dots, x_n)$ et distributions marginales respectives $f_1(x_1), f_2(x_2), \dots, f_n(x_n)$. On dit que X_1, X_2, \dots, X_n sont mutuellement indépendants ssi : $\forall x_i \in \mathbf{R}, \quad f(x_1, x_2, \dots, x_n) = f_1(x_1) \times f_2(x_2) \times \dots \times f_n(x_n)$	
8. Esperance d'une V.A. X de dist. $f(x)$	$\mu = E(X) = \sum_x x f(x)$	$\mu = E(X) = \int_{-\infty}^{+\infty} x f(x) dx$
9. Esperance suite à un changement d'une variable aléatoire à une nouvelle.	Soit une variable aléatoire $X: \Omega \rightarrow F$ de distribution $f(x)$. Soit une nouvelle variable aléatoire $Y = g(X): F \rightarrow R$. On définit l'esperance par : Si $\sum_{x \in F} g(x) f(x) < +\infty$, alors: $\mu_{g(X)} = E(g(X)) = \sum_{x \in F} g(x) f(x)$ Si $\int_{-\infty}^{+\infty} g(x) f(x) dx < +\infty$, alors: $\mu_{g(X)} = E(g(X)) = \int_{-\infty}^{+\infty} g(x) f(x) dx$	

10. Esperance suite à un changement de 2 V.A. en une nouvelle.	Soit deux variables aléatoires X et Y de distribution jointe $f(x, y)$. Soit une nouvelle variable aléatoire $Z = g(X, Y)$. On définit l'espérance par :	
	Si $\sum_x \sum_y g(x, y) f(x, y) < \infty$: $\mu_{g(X, Y)} = E(g(X, Y))$ $= \sum_x \sum_y g(x, y) f(x, y)$	Si $\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x) f(x) dx dy < \infty$ $\mu_{g(X, Y)} = E(g(X, Y))$ $= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} g(x) f(x) dx dy$
11. Propriétés de l'espérance	<ul style="list-style-type: none"> • $E(aX + b) = aE(X) + b$. • $E(X \pm Y) = E(X) \pm E(Y)$ • Si X et Y sont indépendantes alors $E(X \times Y) = E(X) \times E(Y)$ (Attention : la réciproque est fausse) 	
12. Variance σ^2 et écart-type $\sigma = \sqrt{\sigma^2}$	Soit X^2 intégrable c'est à dire : $\sum x^2 f(x) < \infty$. $\sigma^2 = E((X - \mu)^2)$ $= \sum_x (x - \mu)^2 f(x)$	Soit X^2 intégrable c'est à dire : $\int_{-\infty}^{+\infty} x^2 f(x) < \infty$. $\sigma^2 = E((X - \mu)^2)$ $= \int_{-\infty}^{+\infty} (x - \mu)^2 f(x) dx$
13. Propriété de la variance	<ul style="list-style-type: none"> • $\sigma^2 = E(X^2) - E(X)^2$ • $\sigma_{aX+b}^2 = a^2 \sigma_X^2$ 	
14. Covariance	Soit deux variables aléatoires X et Y de distribution jointe $f(x, y)$.	
	$\sigma_{XY} = E((X - \mu_X)(Y - \mu_Y)) = \sum_x \sum_y (x - \mu_X)(y - \mu_Y) f(x, y)$	$\sigma_{XY} = E((X - \mu_X)(Y - \mu_Y)) = \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} (x - \mu_X)(y - \mu_Y) f(x, y) dx dy$
15. Propriétés de la covariance	<ul style="list-style-type: none"> • $\sigma_{XY} = E(XY) - E(X)E(Y)$ • Si X et Y sont indépendantes alors $\sigma_{XY} = 0$ (Attention : la réciproque est fausse) 	
16. Relation entre variance et covariance	Soit deux variables aléatoires X et Y de distribution jointe $f(x, y)$. $\sigma_{aX+bY}^2 = a^2 \sigma_X^2 + b^2 \sigma_Y^2 + 2ab \sigma_{XY}$	
17. Propriété de la relation	Si X et Y sont indépendantes alors $\sigma_{aX+bY}^2 = a^2 \sigma_X^2 + b^2 \sigma_Y^2$ (Attention : la réciproque est fausse)	
18. Coefficient de corrélation linéaire r	Soit deux variables aléatoires X et Y . $r = \frac{\sigma_{XY}}{\sigma_X \sigma_Y}$	

Chapitre 4 : distributions discrètes et distributions continues

Soient X et X_i des variables aléatoires ; et soient x et x_i les éléments de l'univers des possibles.

On étudie par la suite différentes distributions de probabilité de X , c'est à dire $P(X = x)$ ou des X_i c'est à dire $P(X_1 = x_1 ; X_2 = x_2 ; \dots ; X_k = x_k)$.

Ce chapitre sera divisé en 2 grandes parties : distributions discrètes et distributions continues.

Afin de rendre les résultats plus clairs, ils seront placés dans des tableaux.

1. Distributions discrètes

Afin d'éliminer toute ambiguïté et toute confusion, les différents cas sont munis d'exemples. On étudiera ainsi le même exemple pour les différents cas pour rendre la distinction entre ces derniers plus facile. L'exemple sera simple et classique, celui de tirer des boules contenues dans une urne ou un sac

Distributions discrètes :	Définitions / espérance et variance:	Conditions d'utilisation et explications :	Exemple courant :
1. Uniforme	<p>Soit k éléments $x_1, x_2 \dots x_k$ $\forall x \in \{x_i / i \in \llbracket 1; k \rrbracket\}$:</p> $f(x; k) = P(X = x) = \frac{1}{k}$ <hr/> $\mu = \frac{\sum_{i=1}^k x_i}{k} \text{ et } \sigma^2 = \frac{\sum_{i=1}^k (x_i - \mu)^2}{k}$	<ul style="list-style-type: none"> • <u>1 seul essai</u> est réalisé • La probabilité est la même pour tous les résultats. 	Soit un sac contenant 2 boules noires, 2 blanches et 2 bleues. La probabilité de tirer une boule noire est 1/3.
2. Bernoulli de paramètre p	<p>$\forall x \in \{0; 1\}$ $b(x; p) = p^x (1 - p)^{1-x}$</p> <p>Avec $P(X = 0) = 1 - p$ et $P(X = 1) = p$</p> <hr/> $\mu = p \text{ et } \sigma^2 = p(1 - p)$	<ul style="list-style-type: none"> • <u>1 seul essai</u> est réalisé • <u>Deux valeurs de X</u> : soit 0 (échec) soit 1 (succès) avec des probabilités respectives $1-p$ et p. 	Soit un sac contenant 1/3 de boules noires et 2/3 boules blanches. La probabilité de tirer 1 noire est $b\left(1; \frac{1}{3}\right) = b\left(0; \frac{2}{3}\right)$ qui est donc aussi la probabilité de ne pas tirer 1 blanche.
3. Binomiale	<p>Soit x le nombre de succès obtenue lors de n essais de Bernoulli de paramètre p:</p> <p>$\forall x \in \llbracket 1; n \rrbracket$ avec $n \in \mathbb{N}$:</p> $b(x; n; p) = C_n^x p^x (1 - p)^{n-x}$ <hr/> $\mu = np$ $\sigma^2 = np(1 - p)$	<ul style="list-style-type: none"> • <u>n essais de Bernoulli indépendants avec x succès.</u> • La population est infinie – les différents essais sont indépendants , la probabilité de succès reste constante d'un essai à l'autre. 	Soit un sac contenant 1/3 de boules noires et 2/3 boules blanches. La probabilité de tirer 3 noires sur 7 essais est $b\left(3; 7; \frac{1}{3}\right)$. Rmq : il faut remettre la boule tirée après chaque essai si l'échantillon est fini.
4. Multinomial	Soit une expérience à k résultats possibles d'occurrences respectives x_i , répétée n fois, tq:	<ul style="list-style-type: none"> • <u>C'est une binomial mais avec $k > 2$ résultats possibles pour chaque essai</u> (ce n'est plus un processus de Bernoulli). 	Soit un sac contenant 1/3 de boules noires 2/3 blanches et 0 bleues. Sur 8 essais, la probabilité de tirer 3 noires, 1 blanche et 4 bleues est :

	$P(X_i = x_i) = p_i \text{ et } \sum_{i=1}^k x_i = n$ <p>Avec $i \in \llbracket 1; k \rrbracket$.</p> $f(x_1, x_2, \dots, x_k; n; p_1, p_2, \dots, p_k) = P(X_1 = x_1; \dots; X_k = x_k) = \frac{n!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}$	<ul style="list-style-type: none"> La population est infini – <u>les différents essais sont indépendants</u>, les probabilités de succès restent constantes d'un essai à l'autre. 	$f\left(3, 1, 4; 8; \frac{1}{3}, \frac{2}{3}, 0\right)$ <p>Rmq : il faut remettre la boule tirée après chaque essai si l'échantillon est fini.</p>
5. Hypergéométrique	<p>Soit une population finie de N éléments, dont k d'eux sont des succès. En choisissant un échantillon quelconque de la population contenant n éléments, on définit :</p> <p>$\forall x \in \llbracket 1; k \rrbracket$:</p> $h(x, N, n, k) = \frac{(C_k^x \times C_{N-k}^{n-x})}{C_N^n}$ $\mu = \frac{nk}{N} \text{ et } \sigma^2 = \frac{N-n}{N-1} \times n \times \frac{k}{N} \left(1 - \frac{k}{N}\right)$	<ul style="list-style-type: none"> <u>n essais avec x succès, mais qui ne sont plus indépendants.</u> <u>La population est finie</u> – la probabilité de succès d'un essai influe sur celle de l'essai suivant. <u>C'est donc une binomiale mais avec une population finie.</u> 	<p>Soit un sac contenant en total 9 boules dont 1/3 noires et 2/3 blanches. Quelle est la probabilité de tirer exactement 2 boules blanches parmi 5 tirées?</p> <p>La population est de N=9 ; le nombre de succès total est k = 9x2/3 = 6 ; le nombre de succès cherché est x=2 et le nombre d'essais total est n=5. Donc on a :</p> $h(2, 9, 5, 6)$ <p>Rmq : les tirages se font sans remise sinon on est dans une binomiale.</p>
6. Hypergéométrique multivariée	<p>Soit une population finie de N éléments. Soit k classes de cette population contenant respectivement chacune n_1, n_2, \dots, n_k éléments (avec $\sum_{i=1}^k n_i = N$). En choisissant un échantillon de la population à n éléments, on définit :</p>	<ul style="list-style-type: none"> <u>C'est une Hypergéométrique mais avec k > 2 résultats possibles pour chaque essai .</u> <u>La population est finie</u> – la probabilité de succès d'un essai influe sur celle de l'essai suivant. <u>C'est en quelque terme une multinomiale mais avec une population finie.</u> 	<p>Soit un sac contenant en total 12 boules dont 1/3 noires, 1/4 blanches et les restes bleues. Avec 5 tirées, quelle est la probabilité d'avoir 2 noirs, 1 blanche et 2 bleues ?</p> <p>La population est de N=12 ; le nombre de noires total est $n_1 = 4$; de blanches $n_2 = 3$; de bleues $n_3 = 5$; le nombre de noires cherché est $x_1 = 2$; de blanches est $x_2 = 1$; de bleues</p>

	$\forall x_i \in \llbracket 1; n_i \rrbracket \text{ tq } \sum_{i=1}^k x_i = n$ $f(x_1, x_2, \dots, x_k; N; n; n_1, n_2, \dots, n_k)$ $= \frac{(C_{n_1}^{x_1} \times C_{n_2}^{x_2} \times \dots \times C_{n_k}^{x_k})}{C_N^n}$		<p>est $x_3 = 2$; le nombre d'essais total est $n=5$. Donc on a :</p> $f(2,1,2; 12; 5; 4,3,5)$ <p>Rmq : les tirages se font sans remise sinon on est dans une multinomial.</p>
7. Binomiale négative de paramètre k et p	<p>Soit un processus de Bernoulli de paramètre p. Soit $x \geq k$ le nombre d'essais nécessaire pour avoir k succès.</p> $b^*(x; k; p) = C_{x-1}^{k-1} p^k (1-p)^{x-k}$ $\mu = \frac{k(1-p)}{p}$ $\sigma^2 = \frac{k(1-p)}{p^2}$	<ul style="list-style-type: none"> Il faut que la <u>population soit infinie – les essais soient indépendants – le tirage avec remise.</u> Pour comparer avec la binomiale où on cherche le nombre de succès (variable) en fonction d'un nombre d'essais (fixe), la binomiale négative cherche le nombre d'essais (variable) en fonction d'un nombre de succès (fixe). 	<p>Soit un sac contenant 1/3 de boules noires et 2/3 boules blanches. On cherche à obtenir 5 boules noires. Quelle est la probabilité de les obtenir suite à 9 tirages ?</p> <p>On a $k=5$, $x=9$ et $p=1/3$. Donc c'est $b^*\left(9; 5; \frac{1}{3}\right)$</p>
8. Géométrique de paramètre p	<p>Soit un processus de Bernoulli de paramètre p. Soit x le nombre d'essais nécessaire pour avoir le premier succès.</p> $g(x, p) = p(1-p)^{x-1}$ $\mu = \frac{(1-p)}{p} \text{ et } \sigma^2 = \frac{(1-p)}{p^2}$	<ul style="list-style-type: none"> Il faut que la <u>population soit infinie – ou que les essais soient indépendants.</u> C'est une binomiale négative avec 1 seul succès $k = 1$. 	<p>Soit un sac contenant 1/3 de boules noires et 2/3 boules blanches. Quelle est la probabilité d'obtenir la première boule noire suite à 9 tirages ?</p> <p>On a $x=9$ et $p=1/3$. Donc c'est $g\left(9; \frac{1}{3}\right)$</p>

9. Poisson de paramètres λ et t	<p>Soit t le temps ou l'espace en question. Soit λ la moyenne d'occurrences par unité de temps ou d'espace.</p> <p>On définit : $p(x, \lambda t) = \frac{e^{-\lambda t} (\lambda t)^x}{x!}$</p>	<ul style="list-style-type: none"> • Pas de mémoire c'est à dire : le nombre d'occurrences est indépendant d'un intervalle de temps ou d'une région de l'espace à l'autre. • La probabilité d'avoir une unique occurrence dans un petit(e) intervalle de temps ou région est proportionnelle à la taille de ce(tte) dernier(e), indépendamment des occurrences trouvées ailleurs. • La probabilité d'avoir plus qu'une occurrence dans un petit(e) intervalle de temps ou région est négligeable. 	<p>Rmq : On peut approximer la distribution de Poisson à la binomiale $n \rightarrow \infty$ et $p \rightarrow 0$ et $np = cte$, et donc on obtient $p(x, np)$.</p>
	$\mu = \lambda t$ et $\sigma^2 = \lambda t$		

2. Distributions continues

On considère par la suite une variable aléatoire continue X et x un élément de l'univers des possibles.
Soient de plus μ et σ l'espérance et la variance de la variable aléatoire X .

Rappel : Une densité de probabilité est une fonction $f(x)$ (qui peut aussi dépendre d'autres paramètres) représentant une loi de probabilité sous forme d'intégrale. Elle doit vérifier les conditions suivantes :

$$\forall x \in R, f(x) \geq 0 \quad \text{et} \quad \int_{-\infty}^{\infty} f(x) dx$$

Pour une variable aléatoire X , on définit la probabilité sur un intervalle $[a, b]$ par :

$$P(a < X < b) = P(a \leq X \leq b) = \int_a^b f(x) dx$$

Distribution :	Densité f / Moyenne et Variance :	Remarques importantes :
1. Uniforme	<p>Sur $[a ; b]$ on définit :</p> $f(x, a, b) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{sinon} \end{cases}$ $\mu = \frac{a+b}{2} \quad \text{et} \quad \sigma^2 = \frac{(b-a)^2}{12}$	
2. Normale (en général)	<p>Sur $] -\infty ; +\infty[$ on définit :</p> $f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ <p>μ espérance et σ variance données</p>	<ul style="list-style-type: none"> Calculer à chaque fois une probabilité en intégrant cette fonction serait gênant, pour cela on a pensé à faire des tables de références qui seront plus pratiques. Cependant, il serait impossible de faire les tables pour chaque valeur de μ et σ. La bonne nouvelle est qu'avec un changement de variable, on peut transformer n'importe quelle variable aléatoire X suivant une loi normale, en une variable aléatoire Z d'espérance $\mu = 0$ et d'écart-type $\sigma = 1$, par un changement de variable $Z = \frac{X-\mu}{\sigma}$. Cette dernière suit une loi normale particulière, qu'on désigne par loi normale standard. Théorème : Soient les X_i des variables aléatoires indépendantes suivant des lois normales, ayant des espérances μ_i et des variances σ_i^2, et soient des constantes a_i. Toute combinaison linéaire de ces variables $Z = \sum a_i X_i$ suit aussi une loi normale et a comme espérance $\mu = \sum a_i \mu_i$ et comme variance $\sigma^2 = \sum a_i^2 \sigma_i^2$.
3. Normale standard ou centrée réduite	<p>Sur $] -\infty ; +\infty[$ on définit :</p> $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}$ <p>$\mu = 0$ et $\sigma = 1$</p>	<ul style="list-style-type: none"> Dans les TD, on ne va pas calculer les probabilités en intégrant cette fonction, mais en utilisant l'abaque ou la calculatrice. Théorème : Soit X une variable aléatoire binomiale avec $\mu = np$ et $\sigma^2 = np(1-p)$. Quand $n \rightarrow \infty$, alors $Z = \frac{X-np}{\sqrt{np(1-p)}}$ tend vers une standard normale. En pratique, cela est applicable si $\mu = np > 5$ et $\sigma^2 = np(1-p) > 5$.

4. Gamma de paramètres α et β	<p>Sur $]0; +\infty[$ on définit :</p> $f(x) = \begin{cases} \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$ <p>Avec : $\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} dx$ pour $\alpha > 0$</p> <p>$\mu = \alpha\beta$ et $\sigma^2 = \alpha\beta^2$</p>	
5. Exponentielle de paramètre β	<p>Sur $]0; +\infty[$ on définit :</p> $f(x) = \begin{cases} \frac{1}{\beta} e^{-\frac{x}{\beta}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$ <p>$\mu = \beta$ et $\sigma^2 = \beta^2$</p>	<ul style="list-style-type: none"> La distribution exponentielle est un cas particulier de la distribution gamma avec $\alpha = 1$.
6. Bêta de paramètres a et b	$f(x) = \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} & \text{si } 0 < x < 1 \\ 0 & \text{sinon} \end{cases}$ <p>$\mu = \frac{a}{a+b}$ et $\sigma^2 = \frac{ab}{(a+b)^2(a+b+1)^2}$</p>	
7. Log-normale	<p>Soit une variable aléatoire $Y = \ln(X)$ suivant une distribution normale d'espérance μ et d'écart-type σ. On définit la densité de la variable aléatoire X par :</p> $f(x) = \begin{cases} \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(x)-\mu}{\sigma}\right)^2} & \text{si } x \geq 0 \\ 0 & \text{sinon} \end{cases}$	

	$\mu_X = e^{\mu + \frac{\sigma^2}{2}}$ et $\sigma_X = (e^{\sigma^2} - 1)e^{2\mu + \sigma^2}$	
8. Chi 2 (χ^2) à n degrés de liberté	$f(x) = \begin{cases} \frac{1}{2^{\frac{n}{2}} \Gamma(\frac{n}{2})} x^{\frac{n}{2}-1} e^{-\frac{x}{2}} & \text{si } x > 0 \\ 0 & \text{sinon} \end{cases}$ $\mu = n$ et $\sigma^2 = 2n$	<ul style="list-style-type: none"> • C'est un cas particulier de la fonction Gamma avec $\alpha = \frac{n}{2}$ et $\beta = 2$ • C'est la loi d'une variable aléatoire continue $S = \sum X_i^2$ avec les X_i étant n variables aléatoires normale réduites indépendantes.
9. Student	<p>Sur $] -\infty; +\infty[$ on définit :</p> $f(x) = \frac{1}{\sqrt{n\pi}} \frac{\Gamma(\frac{n+1}{2})}{\Gamma(\frac{n}{2})} \left(1 + \frac{t^2}{n}\right)^{-\frac{k+1}{2}}$	<ul style="list-style-type: none"> • On définit autrement la loi de Student de paramètre n par la variable aléatoire : $T = \frac{Z}{\sqrt{\frac{Y}{n}}}$ <p>Avec Z une variable aléatoire normale standard , Y suivant une loi χ^2 à n degrés de liberté, et tel que Z et Y sont indépendantes.</p>

Chapitre 5 : Transformations de Variables

1. Moments et fonctions caractéristiques de variables aléatoires

1.1. Moment d'une variable aléatoire

Définition : Soit X une variable aléatoire. On définit le r^{eme} moment de X par :

$$\mu_r = E(X^r) = \begin{cases} \sum_x x^r f(x) & \text{si } X \text{ est discrète} \\ \int_{-\infty}^{+\infty} x^r f(x) dx & \text{so } X \text{ est continue} \end{cases}$$

Remarque : L'espérance $\mu = \mu_1$ et la variance $\sigma^2 = \mu_2 - \mu_1^2$

1.2. Fonction caractéristique

Définition : Soit X une variable aléatoire. On définit sa fonction caractéristique $\Phi_X(t)$ par :

$$\Phi_X(t) = E(e^{itX})$$

Remarque : Comme son nom l'indique, cette fonction est une caractéristique de la loi, autrement dit : si $\forall t, \Phi_X(t) = \Phi_Y(t)$ alors X et Y ont la même distribution de probabilité.

Propriétés :

- $\Phi_{X+a}(t) = e^{iat} \Phi_X(t)$
- $\Phi_{aX}(t) = \Phi_X(at)$
- Soient n variables aléatoires indépendantes X_1, X_2, \dots, X_n de fonctions caractéristiques respectives $\Phi_{X_1}(t), \Phi_{X_2}(t), \dots, \Phi_{X_n}(t)$, et soit $Y = \sum X_n$. On a alors :

$$\Phi_Y(t) = \Phi_{X_1}(t) \times \Phi_{X_2}(t) \times \dots \times \Phi_{X_n}(t)$$

Quelques exemples de fonctions caractéristiques des lois étudiées dans le chapitre 4 :

Lois discrètes	Bernoulli	Binomiale	Géométrique	Poisson
$\Phi_X(t)$	$(1-p) + pe^{it}$	$((1-p) + pe^{it})^n$	$\frac{pe^{it}}{1 - (1-p)e^{it}}$	$e^{\lambda(e^{it}-1)}$
Lois continues	Uniforme	Exponentielle	Normale	
$\Phi_X(t)$	$\frac{\sin\left(\frac{b-a}{2}t\right)}{\frac{b-a}{2}t} e^{it\frac{(a+b)}{2}}$	$\frac{\lambda}{\lambda - it}$	$e^{it\mu - \frac{\sigma^2 t^2}{2}}$	

1.3. Relation entre moments et fonctions caractéristiques

Théorème : Soit X une variable aléatoire de fonction caractéristique $\Phi_X(t)$. On a la relation :

$$\mu_r = i^{-r} \left(\frac{d^r}{dt^r} \Phi_X(t) \right)_{t=0}$$

Il en découle que si $\mu_2 = E(X^2)$ est finie, on a le développement limité en 0 suivant :

$$\Phi_X(t) = 1 + it\mu_1 - \frac{t^2}{2}\mu_2 + o(t^2) = 1 + itE(X) - \frac{t^2}{2}E(X^2) + o(t^2)$$

2. Transformations de variables

2.1. Cas d'une seule variable aléatoire

Théorème 1 :

Soit X une variable aléatoire de distribution de probabilité $f(x)$. Soit $Y = u(X)$ une autre variable aléatoire qui est une transformation bijective entre les valeurs de X et Y , de distribution de probabilité $g(y)$. Soit $w(y)$ la fonction réciproque de $u(x)$ (c'est à dire $y = u(x) \leftrightarrow x = w(y)$). Alors on a :

- Dans le cas où X et Y sont discrètes : $g(y) = f(w(y)) = f(x)$
- Dans le cas où X et Y sont continues : $g(y) = f(w(y)) |J|$
Avec $J = w'(y)$ étant le Jacobien de la transformation en question.

2.2. Cas de deux variables aléatoires

Théorème 2 :

Soient X_1 et X_2 deux variables aléatoires de distribution de probabilité jointes $f(x_1, x_2)$. Soient $Y_1 = u_1(X_1, X_2)$ et $Y_2 = u_2(X_1, X_2)$ deux autres variables aléatoires qui sont des transformations bijectives entre les valeurs de (X_1, X_2) et (Y_1, Y_2) , de distribution de probabilité $g(y_1, y_2)$. Soient $w_1(y_1, y_2)$ et $w_2(y_1, y_2)$ les fonctions réciproques de $u_1(x_1, x_2)$ et $u_2(x_1, x_2)$. Alors on a :

- Dans le cas où X_1, X_2 et Y_1, Y_2 sont discrètes :
 $g(y_1, y_2) = f(w_1(y_1, y_2), w_2(y_1, y_2)) = f(x_1, x_2)$
- Dans le cas où X_1, X_2 et Y_1, Y_2 sont continues :
 $g(y_1, y_2) = f(w_1(y_1, y_2), w_2(y_1, y_2)) |J|$

Avec $J = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}$ étant le Jacobien.

Chapitre 7 : Echantillonnage et estimation

1. Vocabulaire statistique

1.1. Observation :

Valeur d'une variable aléatoire X ayant une certaine distribution ou loi de probabilité $f(x)$.

1.2. Population :

Ensemble des observations auxquelles on s'intéresse dans une expérience statistique.

On parle par la suite de « population $f(x)$ » comme étant une population dont les observations sont des valeurs d'une variable aléatoire X ayant une distribution $f(x)$.

1.3. Echantillon :

Sous-ensemble d'une population.

1.4. Taille d'une population (ou échantillon) :

Nombre d'observations (fini ou infini) constituant une population (ou échantillon) .

1.5. Echantillonnage :

Ensemble d'opérations permettant de sélectionner d'une certaine façon organisée les éléments de l'échantillon.

Remarque : On distingue généralement deux catégories d'échantillonnage : probabiliste (appelé aussi aléatoire) et non probabiliste. Dans la première, on compte un grand nombre de méthodes, dont on s'intéresse dans le cadre de ce cours à une seule parmi elles : « l'échantillonnage aléatoire simple »

1.6. Echantillonnage aléatoire simple :

Un échantillonnage est aléatoire si tous les individus de la population ont la même chance de faire partie de l'échantillon. Il est simple si les prélèvements des individus sont réalisés indépendamment les uns des autres.

En d'autres termes : Soit une population $f(x)$. Un échantillon aléatoire simple de taille n de cette population est une suite de n variables aléatoires X_1, X_2, \dots, X_n indépendantes ayant même distribution de probabilité $f(x)$.

Remarque : Dans toute la suite du chapitre, chaque fois qu'on fait référence à une population, un échantillon ou une variable aléatoire, il faut les considérer comme définis ici.

1.7. Statistique :

Toute application définie uniquement à partir de variables aléatoires d'un échantillon aléatoire. **Conséquence directe** : une statistique est une variable aléatoire !

(Voir des exemples de statistique dans le paragraphe 2.)

2. Exemples de statistiques

On considère par la suite un échantillon aléatoire de taille n , représenté par les variables aléatoires X_1, X_2, \dots, X_n . On lui définit quelques statistiques par la suite :

2.1. Moyenne : $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$

2.2. Médiane :

$$\tilde{X} = \begin{cases} X_m \text{ tel que } 50\% \text{ des } X_i \leq X_m \text{ et } 50\% \text{ des } X_i \geq X_m \text{ si } n \text{ est impair} \\ \frac{X_m + X_k}{2} \text{ tel que } 50\% \text{ des } X_i \leq X_m \text{ et } X_k \text{ et } 50\% \text{ des } X_i \geq X_m \text{ et } X_k \text{ si } n \text{ est pair} \end{cases}$$

2.3. Mode : Valeur la plus fréquente (peut ne pas exister ou ne pas être unique)

2.4. Etendue : intervalle entre la plus petite et la plus grande valeur.

2.5. Variance empirique : $\sigma^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n}$

2.6. Variance empirique modifiée : $S^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1}$

2.7. Chi-2 : $\chi^2 = \frac{(n-1)S^2}{\sigma^2}$

2.8. Proportion : $\hat{P} = \frac{X}{n}$ avec X le nombre de succès dans n essais

3. Distributions d'échantillonnage

3.1. Définition

Comme on avait déjà signaler dans le paragraphe 1.7. , une statistique est une variable aléatoire et donc on peut lui définir une distribution de probabilité, appelée spécifiquement densité d'échantillonnage.

Remarque : on peut appliquer à une statistique tout ce qui est applicable pour une variable aléatoire vu dans le chapitre 3, par exemple l'espérance, la variance ...

3.2. Distribution de s^2

Soit s^2 la variance d'un échantillon aléatoire de taille n pris dans une population normale de variance σ^2 alors la statistique χ^2 a une distribution χ^2 avec $n-1$ degrés de liberté, avec :

$$\chi^2 = \frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$$

3.3. Distribution de la moyenne

Remarque 1 : Soit X_1, X_2, \dots, X_n un échantillon de taille n d'une population de distribution normale d'espérance μ et de variance σ^2 ; et soit \bar{X} sa moyenne. On a : $\mu_{\bar{X}} = \mu$ et $\sigma_{\bar{X}}^2 = \frac{\sigma^2}{n}$.

Remarque 2 : On ne va pas chercher la distribution exacte de la moyenne, mais une approche.

Theoreme : Soit X_1, X_2, \dots, X_n un échantillon de taille n d'une population de distribution quelconque d'espérance μ et de variance σ^2 (connue) ; et soit \bar{X} sa moyenne.

Plus n est grand, plus la distribution de \bar{X} tend vers une normale $N\left(\mu, \frac{\sigma^2}{n}\right)$.

D'où il en découle que plus n est grand, plus la distribution de $Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ tend vers une normale centrée réduite.

Note : En pratique, pour $n \geq 30$, on a une bonne approximation quelque soit la population parente. Pour $n < 30$ il faut que la population parente soit normale.

3.4. Distribution de Student

Theoreme : Soit X_1, X_2, \dots, X_n un échantillon de taille n d'une population de distribution normale d'espérance μ et de variance S^2 (σ^2 étant inconnue) ; et soit \bar{X} sa moyenne.

$T = \frac{\bar{X} - \mu}{\frac{S}{\sqrt{n}}}$ à une distribution de Student à $n - 1$ degrés de liberté et tend vers une normale réduite centrée si n devient grand.

3.5. Distribution F

Soit U et V deux variables aleatoires independant ayant pour distribution respectives chi-2

de γ_1 et γ_2 degrés de libertés. On définit : $F = \frac{\frac{U}{\gamma_1}}{\frac{V}{\gamma_2}}$.

Si s_1^2 et s_2^2 sont les variances de deux variables aleatoires independantes de tailles n_1 et n_2 pris dans 2 populations de variance σ_1^2 et σ_2^2 alors $F = \frac{\frac{s_1^2}{\sigma_1^2}}{\frac{s_2^2}{\sigma_2^2}}$ a une distribution F de $n_1 - 1$ et $n_2 - 1$ degrés de liberté.

3.6. Récapitulatif

A retenir pour la résolution des exercices :

Population parente	σ connue	$n < 30$	$n \geq 30$
Normale	Oui	Z	Z
	Non	T	T
Non normale	Oui	-	Z
	Non	-	Z

4. Estimations

En statistique comme en probabilité, le hasard intervient fortement. Mais en probabilité, on suppose la loi connue précisément (exemple loi binomiale de paramètres n et p) et on cherche à donner les caractéristiques de la variable qui suit cette loi. L'objectif de la statistique est le contraire : à partir de la connaissance de la variable , que peut-on dire de la loi de cette variable ?

On suppose par la suite une variable aléatoire X de distribution $f(x, \theta)$ qui dépend du paramètre θ inconnu. A partir d'un échantillon de X , on va essayer de déterminer le mieux possible la vraie valeur de θ par une des différentes méthodes d'estimations qui seront présentés dans la suite de ce paragraphe.

4.1. Estimation ponctuelle

Cette méthode consiste à calculer une valeur $\hat{\theta}$ qui soit proche de la vraie valeur θ . Cependant, θ peut avoir plusieurs estimations possibles (exemple : à la moyenne μ peut-être associer la moyenne \bar{X} , la médiane \tilde{X} ou le mode). Afin de choisir la meilleure estimation, l'estimateur $\hat{\theta}$ doit vérifier certaines conditions dont on cite deux :

- L'estimateur doit être sans biais du paramètre θ c'est à dire $E(\hat{\theta}) = \theta$.
- L'estimateur doit être le plus efficace (parmi les estimateurs sans biais ayant passé le premier test) c'est à dire il doit avoir la plus petite variance .

4.2. Estimation par intervalle de confiance

Cette méthode consiste à placer θ dans un intervalle avec une très grande probabilité qu'il s'y trouve réellement dans ce dernier. On verra par la suite quelques exemples d'intervalles de confiance pour quelques statistiques déjà définis précédemment.

- **Estimation de la moyenne :**

Soit \bar{x} la moyenne d'un échantillon aléatoire de taille n . L'intervalle de confiance à $(1 - \alpha) \times 100$ % de la moyenne μ est défini par :

- $\bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}} < \mu < \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$ si σ de la population est connue ;
- $\bar{x} - t_{\alpha/2} \frac{s}{\sqrt{n}} < \mu < \bar{x} + t_{\alpha/2} \frac{s}{\sqrt{n}}$ si σ de la population est inconnue.

Avec $z_{\alpha/2}$ (resp. $t_{\alpha/2}$) est la valeur z (resp. t à $n - 1$ degrés de liberté) laissant une surface $\alpha/2$ à droite dans la courbe de la normale standard (resp. Student).

- **Estimation de la différence entre deux moyennes :**

Soit \bar{x}_1 et \bar{x}_2 les moyennes de deux échantillons aléatoires de taille respective n_1 et n_2 . L'intervalle de confiance à $(1 - \alpha) \times 100$ % de la différence des moyennes μ est défini par :

$$- (\bar{x}_1 - \bar{x}_2) - z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + z_{\alpha/2} \sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

si σ_1 et σ_2 des populations sont connues ;

Avec $z_{\alpha/2}$ est la valeur z laissant une surface $\alpha/2$ à droite dans la courbe de la normale standard .

$$- (\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

si $\sigma_1 = \sigma_2$ des populations est inconnue ;

Avec $t_{\alpha/2}$ est la valeur t à $n_1 + n_2 - 2$ degrés de liberté laissant une surface $\alpha/2$ à droite dans la courbe Student. Et $s_p^2 = \frac{(n_1-1)S_1^2 + (n_2-1)S_2^2}{(n_1+n_2-2)}$

$$- (\bar{x}_1 - \bar{x}_2) - t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}} < \mu_1 - \mu_2 < (\bar{x}_1 - \bar{x}_2) + t_{\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

si $\sigma_1 \neq \sigma_2$ des populations sont inconnues.

Avec $t_{\alpha/2}$ est la valeur t à $\frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$ degrés de liberté laissant une surface $\alpha/2$ à droite dans

la courbe Student.

- **Estimation de la variance :**

Soit s^2 la variance d'un échantillon aléatoire de taille n . L'intervalle de confiance à $(1 - \alpha) \times 100$ % de la variance σ^2 est défini par :

$$\frac{(n-1)S^2}{\chi_{\alpha/2}^2} < \sigma^2 < \frac{(n-1)S^2}{\chi_{1-\alpha/2}^2}$$

Avec $\chi_{1-\alpha/2}^2$ et $\chi_{\alpha/2}^2$ sont les valeurs de la distribution chi-2 à $n - 1$ degrés de liberté, laissant respectivement les surfaces $1 - \frac{\alpha}{2}$ et $\frac{\alpha}{2}$ à leur droite.

- **Estimation d'une proportion :**

Soit \hat{p} la proportion de succès dans un échantillon aléatoire de taille n . L'intervalle de confiance à $(1 - \alpha) \times 100$ % du paramètre p de la loi binomiale est défini par :

$$\hat{p} - z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} < p < \hat{p} + z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

Avec $z_{\alpha/2}$ est la valeur z laissant une surface $\alpha/2$ à droite dans la courbe de la normale standard.

Remarque : cela est valable si $n\hat{p} \geq 5$ et $n(1 - \hat{p}) \geq 5$.

Remarques sur les intervalles de confiance :

Comme on le sait, généralement l'estimation ponctuelle n'est pas exactement égale au paramètre recherché : un certain décalage existe qu'on va appeler marge d'erreur. Parfois on cherche à limiter notre erreur, cela va dépendre évidemment de la taille n de l'échantillon. On définit alors les marges d'erreurs et les tailles suivantes :

- Si \bar{x} est utilisée comme estimation ponctuelle de μ (σ connue), on peut affirmer qu'on est à $(1 - \alpha) \times 100$ % confiant que l'erreur ne dépassera pas $z_{\alpha/2} \frac{\sigma}{\sqrt{n}}$.

D'où pour qu'on soit à $(1 - \alpha) \times 100$ % confiant que l'erreur ne dépassera pas une certaine valeur e , il faut que $n = \left(z_{\alpha/2} \frac{\sigma}{e}\right)^2$ arrondi au plus proche entier.

- Si \bar{x} est utilisée comme estimation ponctuelle de μ (σ inconnue), on peut affirmer qu'on est à $(1 - \alpha) \times 100$ % confiant que l'erreur ne dépassera pas $t_{\alpha/2} \frac{s}{\sqrt{n}}$.

D'où pour qu'on soit à $(1 - \alpha) \times 100$ % confiant que l'erreur ne dépassera pas une certaine valeur e , il faut que $n = \left(t_{\alpha/2} \frac{s}{e}\right)^2$ arrondi au plus proche entier.

- Si \hat{p} est utilisée comme estimation ponctuelle de p , on peut affirmer qu'on est à $(1 - \alpha) \times 100$ % confiant que l'erreur ne dépassera pas $z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$.

D'où pour qu'on soit à $(1 - \alpha) \times 100$ % confiant que l'erreur ne dépassera pas une certaine valeur e , il faut que $n = \frac{z_{\alpha/2}^2}{e^2} \hat{p}(1 - \hat{p})$ arrondi au plus proche entier.

4.3. Estimation par maximum de vraisemblance

Soit toujours la variable aléatoire X de distribution $f(x, \theta)$. Soient n observations indépendantes x_1, x_2, \dots, x_n . L'estimateur de maximum de vraisemblance $\hat{\theta}$ du paramètre θ est la valeur qui maximise la fonction de vraisemblance définie par :

$$L(x_1, x_2, \dots, x_n; \theta) = \prod_{i=1}^n f(x_i; \theta)$$

D'où en pratique on doit calculer $\frac{dL}{d\theta} = 0$ ou $\frac{d \ln(L)}{d\theta} = 0$.

4.4. Estimation par la méthode des moments

Cette méthode consiste à égaliser les moments de la distribution de probabilité de la population dont on cherche le paramètre θ avec les moments empiriques de l'échantillon calculés à partir de ce dernier.

En d'autres termes il faut résoudre :

$$\mu_r(\text{population}) = \mu_r(\text{échantillon}) \quad \forall r \in \mathbb{N}$$

Avec pour rappeler $\mu_r = E(X^r)$.

On prend autant de valeurs de r qu'il y en a de paramètres inconnus (c'est à dire si θ était un vecteur d'ordre k , il faut prendre $r = 1, 2, \dots, k$ et donc résoudre k équations à k inconnues).

Chapitre 8 : Tests d'hypothèses

1. Vocabulaire

1.1. Hypothèse statistique :

C'est une affirmation (qui peut être vraie ou fausse) concernant une ou des caractéristiques (valeurs des paramètres, forme de la distribution des observations) d'une population.

1.2. Hypothèse nulle H_0 et hypothèse alternative H_1 :

L'hypothèse nulle est l'hypothèse selon laquelle on fixe à priori un paramètre de la population à une valeur particulière.

L'hypothèse alternative est n'importe quelle autre hypothèse qui diffère de l'hypothèse nulle.

1.3. Test d'hypothèse (ou test statistique) :

C'est une démarche qui a pour but de fournir une règle de décision permettant, sur la base de résultats d'échantillon, de faire un choix entre deux hypothèses statistiques (voir la démarche à suivre dans le paragraphe 2.).

Ainsi suite au test, on peut soit « accepter » H_0 , soit la rejeter en adoptant H_1 .

Remarque : Pour être plus clair, on ne peut jamais accepter H_0 au sens de véracité, c'est à dire on ne peut pas affirmer par le test que H_0 est vrai ; on peut juste ne pas la réfuter tant que H_1 n'a pas été acceptée. On parle donc de non-rejet et non pas d'acceptation de H_0 .

1.4. Erreurs :

On parle d'erreur de type I suite à un rejet de l'hypothèse nulle qui est vraie.

On parle d'erreur de type II suite à un non-rejet de l'hypothèse nulle qui est fausse.

1.5. Niveau de signifiante (ou seuil de signification) α :

C'est la probabilité de commettre une erreur de type I. D'où on peut affirmer que la probabilité du non-rejet de H_0 qui est vraie est de $1 - \alpha$.

Remarque : La probabilité de commettre une erreur de type II est notée β . D'où on peut affirmer que la probabilité du rejet H_0 qui est fausse est de $1 - \beta$.

Cependant β n'est pas l'objectif de notre cour, on va juste se concentrer sur α .

1.6. Valeur critique :

Valeur limite à partir de laquelle on peut rejeter H_0 avec un niveau de signifiante α .

1.7. P-valeur :

C'est le plus petit seuil de signifiante pour lequel H_0 reste encore non-rejeté.

Ainsi, pour un niveau de signifiante α :

Si $p < \alpha$ alors H_0 est rejetée en faveur de H_1 ; Si $p > \alpha$ alors H_1 est rejetée en faveur de H_0 .

1.8. Statistique de test (notée ST ou T) :

C'est n'importe quelle statistique calculée à partir d'un échantillon prélevé d'une population.

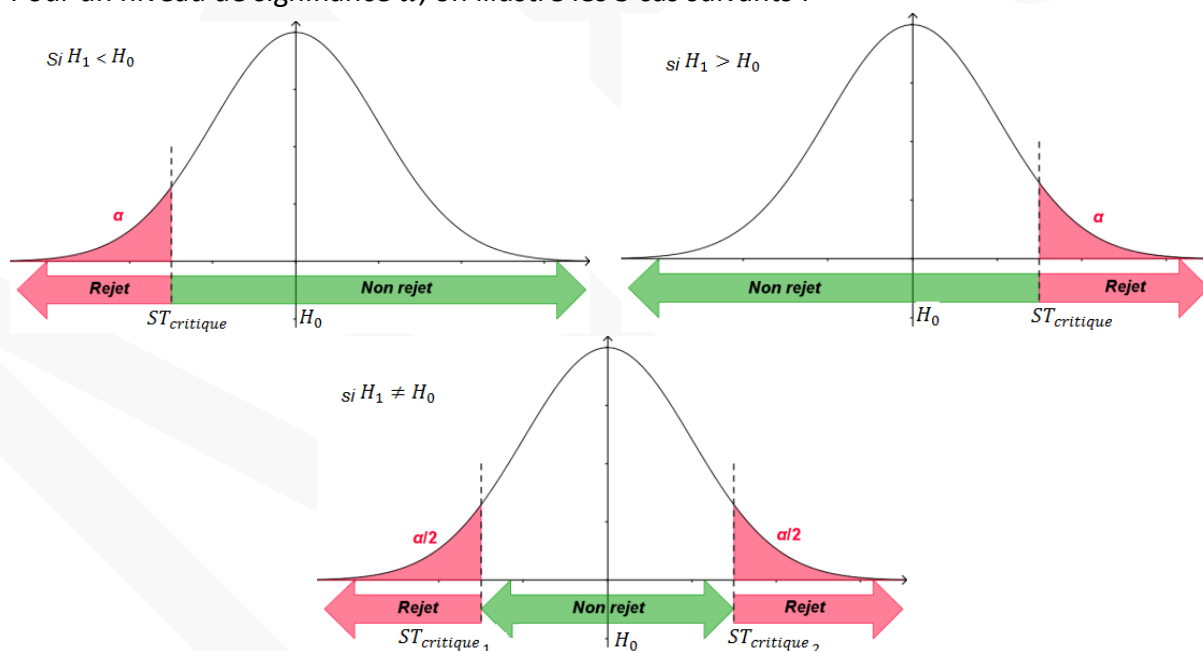
2. Démarche à suivre lors d'un test d'hypothèse sur des paramètres

- 1) Fixer l'hypothèse nulle $H_0 = b$.
- 2) Poser l'hypothèse alternative $H_1 < b$ ou $H_1 > b$ ou $H_1 \neq b$.
- 3) Choisir la statistique (la loi) convenable au problème.
- 4) Déterminer la valeur critique $ST_{critique}$
- 5) Calculer la valeur correspondante à l'échantillon en question $ST_{calculée}$.
- 6) Comparer $ST_{critique}$ et $ST_{calculée}$ puis conclure sur le rejet ou le non-rejet de H_0 .

Remarque : Comme déjà défini dans le paragraphe 1., l'hypothèse alternative H_1 est n'importe quelle autre hypothèse qui diffère de l'hypothèse nulle H_0 . Ainsi, on distingue deux types de tests :

- Le test à queue unique où l'hypothèse alternative est une inégalité $H_1 < H_0$ ou $H_1 > H_0$.
- Le test à double queue où l'hypothèse alternative est une différence $H_1 \neq H_0$.

Pour un niveau de signifiante α , on illustre les 3 cas suivants :



Ainsi si $ST_{calculée}$ se trouve dans la zone rouge, H_0 est rejeté.

3. Tests usuels sur des paramètres

On va étudier quelques tests très souvent employés, concernant des échantillons obéissant à des distributions déjà étudiées dans les chapitres précédents, dont on test certains paramètres.

Test sur :	Statistique de test ST :	Régions critiques pour un niveau de confiance α :	Conditions d'utilisation :
1. Moyenne $H_0:$ $\mu = \mu_0$	$Z = \frac{\bar{X} - \mu_0}{\frac{\sigma}{\sqrt{n}}}$	<ul style="list-style-type: none"> • Si $H_1 < H_0$: $z < -z_\alpha$. • Si $H_1 > H_0$: $z > z_\alpha$. • Si $H_1 \neq H_0$: $z < -z_{\alpha/2}$ et $z > z_{\alpha/2}$. 	<ul style="list-style-type: none"> • σ connue. • Population parente normale ou $n \geq 30$.
	$T = \frac{\bar{X} - \mu_0}{\frac{S}{\sqrt{n}}}$	<ul style="list-style-type: none"> • Si $H_1 < H_0$: $t < -t_{\alpha, n-1}$. • Si $H_1 > H_0$: $t > t_{\alpha, n-1}$. • Si $H_1 \neq H_0$: $t < -t_{\alpha/2, n-1}$ et $t > t_{\alpha/2, n-1}$. 	<ul style="list-style-type: none"> • σ inconnue. • Population parente normale.
2. Différence entre deux moyennes $H_0:$ $\mu = \mu_{1_0} - \mu_{2_0}$	$Z = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_{1_0} - \mu_{2_0})}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$	<ul style="list-style-type: none"> • Si $H_1 < H_0$: $z < -z_\alpha$. • Si $H_1 > H_0$: $z > z_\alpha$. • Si $H_1 \neq H_0$: $z < -z_{\alpha/2}$ et $z > z_{\alpha/2}$. 	<ul style="list-style-type: none"> • σ_1 et σ_2 connues. • Populations parentes normales ou n_1 et $n_2 \geq 30$.
	$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_{1_0} - \mu_{2_0})}{s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$ Avec $s_p^2 = \frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{(n_1+n_2-2)}$	<ul style="list-style-type: none"> • Si $H_1 < H_0$: $t < -t_{\alpha, n_1+n_2-2}$. • Si $H_1 > H_0$: $t > t_{\alpha, n_1+n_2-2}$. • Si $H_1 \neq H_0$: $t < -t_{\alpha/2, n_1+n_2-2}$ et $t > t_{\alpha/2, n_1+n_2-2}$. 	<ul style="list-style-type: none"> • $\sigma_1 = \sigma_2$ inconnue. • Populations parentes normales
	$T = \frac{(\bar{X}_1 - \bar{X}_2) - (\mu_{1_0} - \mu_{2_0})}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$ Avec $\gamma = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{n_1-1} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{n_2-1}}$ degrés de liberté	<ul style="list-style-type: none"> • Si $H_1 < H_0$: $t < -t_{\alpha, \gamma}$. • Si $H_1 > H_0$: $t > t_{\alpha, \gamma}$. • Si $H_1 \neq H_0$: $t < -t_{\alpha/2, \gamma}$ et $t > t_{\alpha/2, \gamma}$ 	<ul style="list-style-type: none"> • $\sigma_1 \neq \sigma_2$ inconnues. • Populations parentes normales
3. Proportion $H_0:$ $p = p_0$	<ul style="list-style-type: none"> • On peut toujours utiliser directement la loi binomiale. • Si p_0 est très proche de 0 ou 1, on peut utiliser la loi de Poisson de paramètre $\mu_0 = np_0$. • Si p_0 n'est pas très proche de 0 ou 1, on peut utiliser la loi normale de paramètres $\mu_0 = np_0$ et $\sigma_0^2 = np_0(1-p_0)$, soit la standard $Z = \frac{x-np_0}{\sqrt{np_0(1-p_0)}}$. 		
4. Variance $H_0:$ $\sigma^2 = \sigma_0^2$	$\chi^2 = \frac{(n-1)S^2}{\sigma_0^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma_0^2}$	<ul style="list-style-type: none"> • Si $H_1 < H_0$: $\chi^2 < \chi_{1-\alpha}^2$. • Si $H_1 > H_0$: $\chi^2 > \chi_\alpha^2$. • Si $H_1 \neq H_0$: $\chi^2 < \chi_{1-\alpha/2}^2$ et $\chi^2 > \chi_{\alpha/2}^2$. 	<ul style="list-style-type: none"> • Population parente normale.

4. Tests sur d'autres caractéristiques

Tous les tests précédents concernent les paramètres d'un échantillon d'une population $f(x)$ (exemples : μ, p, σ^2 ...). Cependant, il est possible de tester d'autres caractéristiques de l'échantillon comme : l'adéquation de la loi $f(x)$ où on cherche si vraiment $f(x)$ peut correspondre à l'échantillon ; l'indépendance de certains caractères ; l'égalité de plusieurs proportions... Comme ces types de tests sont rarement étudiés dans nos exercices, on ne va pas les détailler dans ce résumé afin de le rendre moins dense. En cas de besoin, consulter le cours du prof. pages 155 à 159.