Notes de Cours MAT1720 - Probabilités

Emulie Chhor

May 19, 2021

Introduction

Le premier cours de probabilité comporte 4 chapitres:

1. Introduction aux Probabilités
2. Variables Aléatoires
3. Espérance de Variables Aléatoires
4. Fonctions Génératrices et Théorème Limite Central

1 Introduction à la Probabilité

Overview

Le premier chapitre porte sur les notions de bases en Probabilité. On introduit la notion de dénombrement, la règle de Bayes et la notion d'indépendance.

Il est essentiel de maitriser ces notions puisque les concepts des autres chapitres seront basées sur celles-ci.

1.1 Définition Classique de Probabilité

La définition classique de probabilité nous dit qu'on peut trouver la probabilité en considérant le ratio entre la cardinalité de notre espace et la cardinalité de l'univers des possibles. En d'autres mots, c'est le ratio entre les évènements favorables et tous les évènements.

1.2 Règles de dénombrement

1. Principe de multiplication
2. Principe d'addition

La principe de multiplication et d'addition nous dit que lorsqu'on doit multiplier les évènements consécutifs ensemble, alors que le principe d'addition nous dit d'additionner la probabilité d'évènements disjoints. On verra plus tard que le principe d'addition pourra se généraliser avec le principe d'inclusion-exclusion.

En ce qui attrait le principe de multiplication, je trouve qu'il est utile de visualiser chaque position comme des petites cases qu'on doit déterminer la probabilité à chaque case.

Problème (Permutations et Combinaisons avec et sans remise)**.** *Souvent, les problèmes ne nous disent pas s'il s'agit d'un problème sans remise ou avec remise ou si l'ordre est important. Le plus important est de rester cohérent avec notre stratégie de dénombrement. Si on décide de compter le nombre de groupes, alors il faut rester avec le nombre de groupes. On ne peut pas calculer le nombre de groupe pour l'un et le nombre de personnes pour l'autre*

1.3 Règles de dénombrement II

1. Permutations
2. Combinaisons
3. Boules et Urnes

1.3.1 Permutations

Les permutations comptent le nombre d'évènements en considérant l'ordre.

$$P(n,r) = \frac{n!}{(n-r)!}$$

avec n: nombre d'éléments, r: nombre de cases

Intuitivement, le numérateur nous dit qu'on a n cases et qu'on doit piger 1 boule sans remise à chaque position, alors que le dénominateur nous dit qu'on a compter trop de cases et qu'on doit retrancher les positions qu'on a compter de trop.

1.3.2 Combinaisons

Les combinaisons comptent le nombre d'évènements sans considérer l'ordre de la pige et est donné par

$$C(n,r) = \frac{n!}{r!(n-r)!}$$

Intuitivement, le nombre de combinaisons est le nombre de permutations dont on a retrancher les groupes équivalents, donné par r!

Problème (Compter le nombre de paires)**.** *Lorsqu'on doit dénombrer de paires, on doit considérer la position de la paire et la sortes. Exemples:*

1. *Souliers: paires + gauche/droite*
2. *Cartes: valeur + enseigne*

1.3.3 Boules et Urnes

Le principe de boules et urnes nous dit qu'on veut placer n boules dans r urnes. Dépendamment des problèmes, on accepte que certaines urnes soients vides ou non. Pour ma part, j'aime représenter les urnes avec des petites boites. Parfois, certains problèmes plus tricky requiert qu'on dissecte le problème en plusieurs étapes. On doit dabord fixer des positions, puis résoudre avec l'astuce de boule et urnes par la suite.

Problème (Code Binaire)**.**

1.4 Règles de dénombrement III

1. Formule de Pascal
2. Binôme de Netwon
3. Coefficients Binomiaux

1.4.1 Formule de Pascal

1.4.2 Binôme de Netwon

La formule du Binome de Newton nous permet de développer des puissances de binôme

$$(x+y)^n = \sum_{k=1}^{n} \binom{(}{n}, x) x^{n-k} y^y$$

Remarque. $(x+y)^n = (x+y)^{n-1}(x+y)$

Intuition sur la formule

En calculant les intégrales de $\int (1-x^2)^n$, Newton a remarqué un parttern intéressant: le numérateur des premiers termes se trouvent de la façon suivante:

1. 1er terme: toujours x
2. Numérateur du 2e coefficient: exposant n
3. Numérateur du 3e coefficient: nombre triangulaire avec k=n-1, $\frac{k(k-1)}{2}$
4. Numérateur du 4e coefficient: nombre pyramidal avec k=n-2 $\frac{k(k+1)(k+2)}{1 \cdot 2 \cdot 3} \implies \frac{(n-2)(n-1)(n)}{1 \cdot 2 \cdot 3}$
5. Numérateur du n-ième coefficient: $\frac{n}{1} \cdot \frac{n-1}{2} \cdot ... \cdot \frac{0}{n}$

https://www.youtube.com/watch?v=sZaXmKB5xSI

$$\sum_{i=0}^{n} \binom{(}{n}, k)2^n, \forall n \geq 0$$

On prend x=1 et y=1 dans la formule de Newton

1.4.3 Coefficients Binomiaux

1.5 Espaces de Probabilité

1. Espace Fondamental

Un espace de probabilité est l'univers des possibles, c-à-d toutes les valeurs que l'évènement X peut prendre. On note 3 axiomes importants, duquel découlent tous les autres:

1. $\mathbb{P}(\Omega) = 1$
2. $\mathbb{P}(A^c) = 1 - \mathbb{P}(A)$

1.5.1 Propriété de l'Espace Fondamental

Propriété 1.5.1 (Propriétés de l'Espace fondamental)**.**

Propriété d'additivité

Continuité croissante/décroissante

Additivité Dénombrables

Inégalité de Boole

Propriété de Monotécité

1.5.2 Principe d'inclusion-exclusion

Le principe d'inclusion-exclusion nous permet de caluler la probabilité de l'union (et de l'intersection) d'évènements qui ne sont pas disjoints.

Pour mieux comprendre, on peut dessiner un diagramme de Venne. Essentiellement, la formule nous dit qu'on doit retrancher les intersections pour éviter de compter le même espace deux fois.

1.6 Probabilité Conditionnelle

1. Définition
2. Formule de Bayes
3. Conditionnement

1.6.1 Définition de la Probabilité Conditionnelle

Intuitivement, la définition de la probabilité conditionnelle nous dit que si on sait qu'une partie de l'espace fondamental s'est déjà produit, alors on n'a plus besoin de le considérer (car ça s'est déjà passé), et on peut le retrancher de la probabilité

1.6.2 Formule de Bayes

La formule de Bayes nous dit nous permet de calculer de la probabilité à postériori. En observant un résultat, on peut faire de l'inférence pour peut-être obtenir plus d'information sur une prieure.

On peut retrouver la formule en faisant l'égalité entre 2 probabilités conditionnelles.

1.7 Indépendance

1. Pairwise Independance
2. Indépendance
3. Conditionnement

On peut interpréter l'indépendance de la façon suivante: connaitre le résultat de X n'affecte pas la probabilité de Y. En d'autres mots, si je pige X avant Y, je ne change pas la probabilité de Y. Connaitre le résultat de X ne me donne pas plus d'information sur Y.

On sait que deux évènements (ou plus) sont indépendant si:

1. Pairwise Independance: la probabilité de chaque paire est égale à la probabilité de leur intersection
2. Independance "for all": la probabilité de l'intersection entre A,B,C est égale à leur produit

Notons qu'on veut déterminer l'indépendance entre 2 variables, car ça nous permet d'additionner des v.a. (prochain chapitre)

2 Variables aléatoires

Overview

On étudie les variables aléatoires parce qu'on n'est pas capable de déterminer la "chance" d'obtenir un résultat précis en un seul essai. Par contre, si on répète l'expérience assez de fois, on est en mesure d'estimer la probabilité d'un évènement

Le chapitre sur les variables aléatoires se découlent en:

1. Variables aléatoires discrètes: valeur discrete
2. Variables aléatoires continues: valeur continue
3. Somme et Produit de Variable
4. Transformation de variable aléatoire
5. V.a. Conditionnelles

Il faut aussi se souvenir qu'une v.a. est définie par la fonction densité, représentant la courbe de sa valeur, et par sa fonction de répartition, la valeur de sa probabilité, équivalent à l'aire sous la courbe.

Notons que les concepts du premier chapitre sont encore valide. Ce chapitre essait de généraliser les résultats du premier chapitres. Ainsi, on sait que l'aire sous la courbe est de 1.

2.1 Variables aléatoires discrètes

2.2 Définition

Une variable aléatoire discrète est une variable qui modélise un évènement dont la valeur est discrète. Ici, on désire déterminer:

1. Probabilité que X prenne une valeur quelconque: $P(X = x)$
2. Probabilité que X prenne une valeur dans un intervalle: $P(X \leq x)$ ou $P(X \geq x) = 1 - P(X \leq x)$ ou $P(n < X < m)$

Ainsi, si on veut déterminer la proba que X prenne la valeur 1 ou 2 ou ..., on n'a qu'à additionner ses probabilités là.

Il existe plusieurs variables aléatoires discrètes, qui modélise des situations différentes, et elles se trouvent avec les techniques de dénombrement vues au chapitre 1.

2.3 Loi de Variables aléatoires discrètes

1. Épreuves de Bernouilli
2. Loi Binomiale
3. Loi Géométrique
4. Loi Hypergéométrique
5. Loi Binomiale Négative
6. Loi de Poisson

7. Processus de Poisson

2.3.1 Épreuves de Bernouilli

Malgré qu'on ne la considère pas comme une loi, une épreuve de Bernouilli réprésente la probabilité d'un succès ou d'un échec pour un seul essai.

$$P(I) = \text{ 1 * proba succès + 0 * proba échec}$$

On verra plus tard qu'on l'appelle aussi une variable indicatrice, et nous aide à calculer les calculer la somme de variables aléatoires.

2.3.2 Loi Binomiale

La Binomiale est une généralisation de la Bernouilli. On veut déterminer la probabilité d'avoir k succès en n épreuves. Une hypothèse de la binomiale est que la probabilité de chaque succès est la même.

La formule de la binomiale est assez intuitive. Si la probabilité de succès est donnée par p, alors on sait que la probabilité d'échec est de (1-p). De plus, on a k succès et n-k échecs. Ainsi, on a

$$p^k(1-p)^{n-k}$$

, qu'on multiplie par la combinaisons, car l'ordre n'importe pas.

Notons que si p est petit et n est grand, on peut modéliser avec une poisson (on verra plus tard pourquoi)

2.3.3 Loi Géométrique

La loi géométrique représente la probabilité d'avoir le premier succès au k-ième essai. Elle suppose aussi que la probabilité de succès est la même à chaque essai (tirage avec remise).

Intuitivement, la formule peut être interpréter de la façon suivante: si on veut que le premier succès soit à la k-ième position, on doit fixer la k-ième position à un succès et les autres à un échec.

$$P(X) = (1-p)^n p$$

Notons qu'on ne multiplie pas par la combinaison parce que ici, l'ordre est important

Pourquoi dit-on que la géométrique est sans mémoire

On dit que la géométrique est sans mémoire, car elle ne dépend pas des résultats précédants. Connaitre l'historique ne nous donne pas plus d'information sur ce qu'il va se passer dans le futur: on n'a besoin que de l'étape précédante

2.3.4 Hypergéométrique

Une distribution hypergéométrique détermine la probabilité d'avoir k succès en n épreuves, mais dans un tirage sans remise.

$$\frac{P(X = x) = \left(^{(a,x)\binom{(}{N}-a,n-x)}\right)}{\binom{(}{N}, n)}$$

Avec

1. n: nombre objets
2. a: nombre de succes
3. X: nombre de succès dans le sample

Le dénominateur représente le nombre de façon de piger n objets et le numérateur représente le nombre de façon d'avoir x succès parmi les a succès possibles x la proba d'avoir n-x échecs

2.3.5 Binomiale Négative

La binomiale négative est la distribution qui nous donne la distribution du nombre d'essais requis pour avoir le r-ième succès avec probabilité p.

Pour obtenir la binomiale négative, on doit d'abord fixer le k-ème succès au x-ième essais.

$$p$$

Puis, il faut qu'il y ait (r-1) succès dans les (x-1) essais (essais avant le k-ième succès). On sait aussi que l'ordre n'importe pas, alors on doit multiplier par la combinaison. (utiliser l'intuition de la binomiale)

$$\binom{x-1, r-1}{p}^{r-1}(1-p)^{(x-1)-(r-1)}$$

On obtient donc

$$P(X = x) = px\binom{x-1, r-1}{p}^{r-1}(1-p)^{(x-1)-(r-1)} = \binom{x-1, r-1}{p}^{r}(1-p)^{(x-r)}$$

2.3.6 Poisson

TODO

La loi de Poisson mesure le nombre d'évènements dans un intervalle donné.

2.3.7 Processus de Poisson

2.4 Variables aléatoires continues

2.5 Définition

Une variable aléatoire continue est définie par sa fonction de densité, qui donne la "valeur" de la probabilité à un point donné, et sa fonction de répartition, qui donne la probabilité que X prenne une valeur dans un intervalle donné. Notons que la probabilité que X prenne une valeur quelconque est de 0, car l'aire sous la courbe d'un point est nulle. Ainsi, on ne veut pas calculer la probabilité que X prenne un valeur quelconque, mais plutôt:

1. Fonction de densité: valeur de X a un point
2. Fonction de Répartition: proba que X prenne une valeur dans l'intervalle $P(X \leq x)$

2.6 Loi de Variables aléatoires continues

Il existe quelques lois continues:

1. Loi Uniforme
2. Loi Exponentielle
3. Loi Normale
4. Loi Log-Normale
5. Loi Gamma
6. Loi Chi-Deux
7. Loi Cauchy

Nous nous focusserons sur les primordiales, c-à-d la loi uniforme, exponentielle, et normale. On peut trouver la fonction de répartition en intégrant la fonction de densité, et, inversement, trouver la fonction de densité en dérivant la fonction de répartition

Comment trouver la fonction de densité de v.a continues?

TODO

2.6.1 Loi Uniforme

La loi uniforme est dont la probabilité est la même partout. TODO

2.6.2 Loi Exponentielle

TODO

2.6.3 Loi Normale

TODO

Remarque. *On peut utiliser la loi normale pour trouver la formule de Stirling, qui approxime la factorielle n!*

Remarque (Loi Log-Normale)**.** *Si on a une loi log normale, on a qu'à poser une variable intermédiaire log(D) et travailler avec l'exponentielle*

2.7 Variables aléatoires simultanées

Les variables aléatoires simultanées sont des variables qui considèrent 2 paramètres (ou plus). Par exemple, on calcule le BMI à partir du poids et de la taille.

On distingue 2 types de v.a. simultanées: discrètes et continues, et chacune d'elles possèdent

1. Fonction de Masse Conjointe: fonction de probabilité
2. Fonction de Masse Marginale: fonction "individuelle"
3. Fonction de Masse Conditionnelle:

2.8 Définitions v.a. simultanées discrètes

2.8.1 Fonction de Masse Conjointe

La fonction de masse conjointe nous donne la probabilité de la variable simultanée. On écrit

$$p(x, y) = P(X = x, Y = y) = P(X = x|Y = y)P(Y = y)$$

Intuitivement, c'est l'intersection de X et de Y, et puisque l'intersection est donnée par la règle de multiplication (voir chapitre 1), on obtient la proba conditionnelle plus haut.

Pourquoi la fonction de masse conjointe peut être interpréter comme l'intersection de X et Y

TODO

2.8.2 Fonction de Masse Marginale

La fonction de masse marginale de X représente l'apport de X et on la trouve en sommant la fonction de masse conjointe p/r au support de l'autre variable

Pourquoi doit-on sommer/intégrer p/r à l'autre variable lorsqu'on calcule la fonction de masse marginale?

TODO

2.8.3 Fonction de masse conditionnelle

C'est le même principe que la probabilité conditionnelle, mais on considère le ratio des proportions.

2.9 Définitions v.a. simultanées continues

1. Fonction de Densité Conjointe
2. Fonction de Densité Marginale
3. Fonction de Densité Conditionnelle

2.10 Variables aléatoires Indépendates

On peut déterminer si X et Y sont indépendantes comme au chapitre 1, c-à-s qu'on compare leur probabilité, qu'on calcule avec l'intégrale.

De savoir que deux variables simultanées sont indépendantes est pratique, car ça nous permet de trouver leur fonction de masse/densité conjointe en les multipliant.

2.10.1 Somme de Variables Aléatoires Indépendantes

TODO

Qu'est-ce qu'une convolution?

TODO

Problème (Trouver la fonction densité de Z=X+Y)**.** *Pour trouver la fonction densité de Z, on doit dabors trouver sa fonction de densité en intégrant p/r à X ou Y, puis on dérive*

Proposition 2.1. *La somme de v.a. normales est la somme de leurs parametres*

2.10.2 Produit de Variables Aléatoires Indépendantes

TODO

subsectionLoi Binormale

2.11 Transformation de Variables Aléatoires

Si on a 2 variables U et V définies par X et Y et on connait que les fonctions densités de X et de Y, il faut passer par une transformation. Par exemple, U=Y+X et V=Y-X.

Pour ce faire, on doit réécrire les fonctions Y et X p/r à U et V. Par exemple, X=U+V et Y=V-U. Par la suite, on doit calculer la jacobienne avec les dérivées partielles. On n'a qu'à appliquer la formule

Pourquoi doit-on multiplier par la jacobienne?

On multiplie par la jacobienne, car puisqu'on pose de nouvelle variable, on change la base du système et le quadrillage change. Ce n'est donc plus des rectangles qu'on sommme, mais bien des parallélogrammes, et on sait que les parallélogrammes est donné par le déterminant.

3 Espérance de Variables Aléatoires

Overview

On peut interpréter l'espérance comme une moyenne pondérée, et la variance comme la distance moyenne entre un point et la moyenne.

1. Espérance et Variance de v.a. discrètes
2. Espérance et Variance de v.a. continues
3. Espérance et Variance de v.a. indépendantes
4. Espérance et Variance de v.a. simultanées

3.1 Linéarité de l'Espérance

Calculer la variance d'une v.a. peut être compliquée, mais on utilise la linéarité de l'espérance pour la calculer. Comme on le sait, la variance est l'écart entre La moyenne et le point. On a donc $V(X) = E((X - \mu)^2)$. On appelle aussi le moment d'ordre 2. En développant, on obtient $V(X) = E(X^2) - E(X)^2$.

3.2 Espérance et Variance de v.a. discrètes

On peut trouver l'espérance et la variance avec les formules générales

1. Espérance: E(X)
2. Moment d'ordre 2: $E(X^2)$
3. Variance: $V(X) = E(X^2) - E(X)^2$

mais on connait l'espérance et la variance des distributions classiques:

1. Binomiale
   - Espérance
   - Variance
2. Géométrique
   - Espérance
   - Variance
3. Poisson

- Espérance
- Variance

3.3 Espérance et Variance de v.a. continues

Tout comme l'espérance et la variance d'une v.a. discrète, on peut interpréter l'espérance d'une v.a. continue comme la moyenne pondérée. Cette fois-ci, on a

1. Espérance: $\int_{-\infty}^{]infty} xp(x)\, d$
2. Moment d'ordre 2:
3. Variance: $V(X) = E(X^2) - E(X)^2$

On connait aussi l'espérance et la variance des distributions continues classiques

1. Uniforme
   - Espérance
   - Variance
2. Exponentielle
   - Espérance
   - Variance
3. Normale
   - Espérance
   - Variance

3.4 Espérance et Variance de v.a. indépendantes

3.5 Espérance et Variance de v.a. simultanées

3.6 Espérance et Variance de v.a. Conditionnelle

4 Fonction Génératrice et Théorème Limite Central

Overview

4.1 Fonction Génératrice

Si deux variables X et Y ont la même fonction génératrice, alors elles ont la même fonction de densité de répartition, et ainsi, elles ont la même espérance et variance.

$$M_X(t) = E(e^{tX})$$

Pour calculer le moment d'ordre 1 et 2, nous permettant de trouver l'espérance et la variance, on doit évaluer la dérivée première et deuxième en zéro.

$$M_X^{(}n)(0) = E(X^n)$$

1. Fonction génératrice Normale
2. Fonction génératrice Gamma
3. Fonction génératrice Poisson
4. Fonction Binomiale et Binomiale Négative

4.1.1 Somme de Fonctions Génératrices

Pour trouver la somme de fonctions génératrices, on multiplie les fonctions génératrices:

$$M_{X_1+...+X_n} = M_{X_1}x...xM_{X_2}$$

4.2 Théorèmes Limites

4.3 Loi des Grands Nombres

4.4 Inégalité de Bienyamé-Tchebychev

L'inégalité de Bienyamé-Tchebychev donne un majorant pour la probabilité. Si on veut trouver le minorant, n'a qu'à utiliser les probabilités totales.

Parfois, le majorant est proche de la vraie valeur, parfois non.

Comment savoir si BT va être précis?

TODO

4.5 Théorème Limite Central

TODO