Soutien pour le cours de probabilités

Partie II: Variables aléatoires discrètes.

Anna Melnykova

Ensimag

2020-2021

...previously on "Soutien pour le cours de probabilités"

- Ω — ensemble d'événtualités (ou ensemble d'issues)
- $\mathcal{F}$ — une collection non vide des événements
- $\mathbb{P}$ — mesure de probabilité
- $(\Omega, \mathcal{F}, \mathbb{P})$ — espace de probabilités (espace probabilisé)

Rappel

Definition (Tribu)

Une tribu $\mathcal{F}$ (ou "**sigma-algebra**" en anglais) sur un espace fondamental Ω est une collection non vide de sous ensembles de Ω vérifiant les deux conditions suivantes :

(i) $\forall A \in \mathcal{F}, \quad A^c \in \mathcal{F}$

(ii) Pour toute suite $A_i \in \mathcal{F},\ i \geq 1$: $\bigcup_{i=1}^{\infty} A_i \in \mathcal{F}$

Definition (Mesure de probabilité)

Une mesure de probabilité sur $(\Omega, \mathcal{F})$, notée $\mathbb{P}$ est une fonction de $\mathcal{F} \to [0,1]$ vérifiant

- $\mathbb{P}(\Omega) = 1$;
- Pour toute suite $A_i \in \mathcal{F},\ i \geq 1$ d'événements disjoints ($A_j \cap A_j = \emptyset$ si $i \neq j$), on a

$$\mathbb{P}\left(\bigcap_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mathbb{P}(A_i)$$

Probabilités conditionelles. Propriétés.

Soient $B_1, B_2, \ldots$ une suite d'événements mutuellement exclusifs $(B_i \cap B_j = \emptyset \quad \forall i \neq j)$ tels que $\forall n \geq 1$, $\mathbb{P}(B_n) \neq 0$ et $\bigcup_{n\geq 1} B_n = \Omega$.
Alors :

- **Formule des probabilités totales** :

$$\forall A \in \mathcal{F},\ \mathbb{P}(A) = \sum_{n\geq 1} \mathbb{P}(A|B_n)\mathbb{P}(B_n)$$

- **Formule de Bayes** :

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A \cap B_i)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum_{n\geq 1} \mathbb{P}(A|B_n)\mathbb{P}(B_n)}$$

Exercice pour rappel

On lance deux dés. Après, on considere une variable S, qui modélise la somme des faces des 2 dés.

1. Quelles sont les éventualités (issues) de lancer de deux dés ? Il y en a combien ? Est-ce qu'ils sont tous equiprobables ?
2. Quelles sont les éventualités de S ? Il y en a combien ? Est-ce qu'ils sont tous equiprobables ?
3. Calculer la probabilité que $S = 2$.
4. Calculer la probabilité que $S > 3$.
5. Calculer la probabilité que $S = 6$ sachant que la nombre sur face d'un de dé est strictement plus petit que 3.

Comment caractériser une probabilité $\mathbb{P}$?

- Ω **est dénombrable (ou fini) :** on peut prendre $\mathcal{F} = \mathcal{P}(\Omega)$, et definir la probabilité pour chaque issue dans Ω. *(Voir exemple avec le lancer de dé)*.
- Ω **est non dénombrable (par exemple,** $\mathbb{R}$**) :** on considère la tribu de Borel $\mathcal{F} = \mathcal{B}(\mathbb{R})$, et on définit une mesure de probabilité sur $\mathbb{R}, \mathcal{B}(\mathbb{R})$ à l'aide d'une fonction $F : \mathbb{R} \to [0, 1]$ telle que $F(x) = \mathbb{P}(]-\infty, x])$, qui s'appelle **fonction de répartition**.

Variables aléatoires réelles

Definition

Toute application mesurable $X : (\Omega, \mathcal{F}) \to (\mathbb{R}, \mathcal{B}(\mathbb{R}))$ s'appelle **variable aléatoire réelle**. X est mesurable si

$$\forall B \in \mathcal{B}(\mathbb{R}) \text{ si } X \in B, \quad X^{-1}(B) \in \mathcal{F}.$$

Autrement dit, pour chaque événement $B \in \mathcal{B}(\mathbb{R})$ (i.e. X prenne une valeur dans l'intervalle B) c'est possible de quantifier la probabilité d'avoir l'événtualite ω telle que $X(\omega) \in \mathcal{B}$. Cette probabilité peut être calculé avec la fonction de répartition, qui caractérise **la loi** X :

$$\mathbb{P}(X \leq x) = \mathbb{P}(\omega \in \Omega : X(\omega) \leq x) = \mathbb{P}(\omega \in \Omega : X(\omega) \in (-\infty, x]) = F(x)$$

Fonction de répartition

Definition

Soit X une variable aléatoire sur $(\Omega, \mathcal{F}, \mathbb{P})$. On note F_X la fonction de $\mathbb{R} \to [0,1]$ définie $\forall x \in \mathbb{R}$ par $F_X := \mathbb{P}(X \leq x)$. La fonction F est appelée **fonction de répartition** de la variable aléatoire X.

Definition

Soit X une variable aléatoire de fonction de répartition F_X. **Le quantile** q_p **(d'ordre** $p \in [0,1]$**)** de X est solution de $F_X(q_p) = p$, i.e.

$$q_p = F_X^{-1}(p).$$

Fonction de répartition : proprietés

1. F_x est croissante au sense large :
   $\forall (x, y) \in \mathbb{R}^2, \quad x < y \Rightarrow F_X(x) \leq F_X(y)$
2. $\lim_{x \to -\infty} F_X(x) = 0 \quad \lim_{x \to +\infty} F_X(x) = 1$
3. F_x est continue à droite : $\forall x \in \mathbb{R}$: $\lim_{h \to 0} F_X(x + h) = F_X(x)$
4. $\lim_{h \to 0} F_X(x - h) = \mathbb{P}(X < x) \leq F_X(x)$

Fonction de répartition : example

- On considere une variable $S : \Omega \to \{2, 3, \ldots, 11, 12\}$, qui modélise la somme des faces des 2 dés.
- Notons que $\forall k \in \{2, 3, \ldots, 11, 12\}$, $\{S = k\} = \{\omega \in \Omega : S(\omega) = k\}$.
- Notons $p_k := \mathbb{P}(S = k)$, on a $\sum_{k=2}^{12} p_k = 1$

La fonction de répartition pour la variable S et défini comme

$$F_S(x) = \mathbb{P}(S \leq x) = \begin{cases} 0 & \text{si } x < 2, \\ \sum_{k=1}^{j} p_k & \text{si } 2 \leq j \leq x < j + 1 \leq 12, \\ 1 & \text{si } x \leq 12. \end{cases}$$

Variable aléatoire discrète

Definition

Une variable aléatoire $X : \Omega \to \mathbb{R}$ est dite **discrète** s'il existe un ensemble fini ou infini dénombrable $E \in \mathbb{R}$ tel que $\mathbb{P}(X \in E) = 1$.

Exemples :

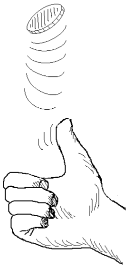
- Pile ou face (lancer d'une pièce)
- Face d'un dé
- Somme de faces de N dés

(Contre-)exemples des ensembles discrets

- Tous les ensembles finis sont discrets
- Les ensembles $\mathbb{N}^d, \mathbb{Z}^d$ sont discrets et dénombrables $\forall d \in \mathbb{N}$
- L'ensemble $\{n \in \mathbb{N}^*, \frac{1}{n}\}$ est discret. On note que cet ensemble est dénombrable et borné : il est contenu dans $[0, 1]$.
- $\mathbb{Q}$ est dénombrable, mais n'est pas discret ! **Par contre, on peut quand même définir une variable qui prenne ses valeurs dans $\mathbb{Q}$!.**

Exemple : lancer d'une pièce

- $X = \mathbb{1}_{\text{la pièce a tombé sur pile}}$
- $\Omega = \{0, 1\}$
- $\mathcal{F} : \{\emptyset, \{0\}, \{1\}, \Omega\}$
- $\mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \frac{1}{2}$
- $F_X(x) = \begin{cases} 0 & \text{si } x < 0, \\ \frac{1}{2} & \text{si } 0 \leq x < 1, \\ 1 & \text{si } x \geq 1 \end{cases}$



Exercice : faire un graphique !

Exemple : lancer d'un dé

- $X =$ nombre sur face d'un dé
- $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $\mathbb{P}(X = k) = \frac{1}{6}$
- $F_X(x) = \begin{cases} 0 & \text{si } x < 1, \\ \frac{k}{6} & \text{si } k \leq x < k+1, \ \forall k \in \{1, \dots, 5\} \end{cases}$

Exercice : faire un graphique !

Fonction de masse

Definition

Soit X une variable aléatoire discrète à valeurs dans un ensemble $E = \{e_1, e_2, \dots\}$ et F_X sa fonction de répartition. On appelle **fonction de masse de** X, la fonction $p_X : E \to [0,1]$ définie par

$$p_X(e_i) := \mathbb{P}(X = e_i) = F_X(e_i) - F_X(e_{i-1}),$$

avec la convention $F_X(e_0) = 0$.

Remarque

Notons que

$$F_X(x) = \sum_{k=1}^{i} p_X(e_k), \quad \forall x \in [e_i, e_{i+1}).$$

Espérance mathématique

Definition

Soit X une variable aléatoire discrète à valeurs dans un ensemble $E = \{e_1, e_2, \dots\}$ et p_X sa fonction de masse. **L'espérance mathématique** de X (appellé aussi moment d'ordre 1) est défini comme

$$\mathbb{E}[X] = \sum_{e_i \in E} e_i p_X(e_i)$$

Exercises

Calculer l'espérance mathèmatique d'une variable X, ou

1. $X = \mathbb{1}_{\text{la pièce tombe sur pile}}$
2. $X = \{\text{nombre sur face d'un dé equilibré}\}$

Espérance mathématique : propriétés

- $\forall a, b \in \mathbb{R}$ et pour les variables aléatoires X et Y

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

- Soit X une variable aléatoire discrète à valeurs dans un ensemble $E = \{e_1, e_2, \dots\}$ et p_X sa fonction de masse. La fonction $g : E \to \mathbb{R}$ est une variable aléatoire discrète et son espérance vaut

$$\sum_{e_i \in E} g(e_i) p_X(e_i).$$

Exemple

Soit X une variable aléatoire discrète à valeurs dans un ensemble $E = \{e_1, e_2, \dots\}$ et p_X sa fonction de masse. Considerons $Y = X^2$.

$$\mathbb{E}[Y] = \mathbb{E}[X^2] = \sum_{e_i \in E} e_i^2 p_X(e_i).$$

Variance

Definition

La variance d'une variable aléatoire X est définie comme la moyenne des carrés des écarts à la moyenne de X :

$$Var[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

1. $\forall c \in \mathbb{R} \quad \mathbb{E}[(X-c)^2] \geq Var[X]$
2. $Var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$ (**Exercice :** prouver !)
3. $Var[aX + b] = Var[aX] = a^2 Var[X]$ (**Exercice :** prouver !)

Variance : propriétés

- Soient X, Y deux variables aléatoires. On a

$$Var[X+Y] = Var[X] + Var[Y] + 2Cov[X,Y], \text{ ou}$$
$$Cov[X,Y] = \mathbb{E}[(X-\mathbb{E}[X])(Y-\mathbb{E}[Y])] = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Si X et Y sont indépendantes, $Cov[X,Y] = 0$ **!**

- Soient $X_1, \ldots, X_n$ n variables aléatoires. On a

$$Var\left(\sum_{k=1}^{n} X_k\right) = \sum_{k=1}^{n} Var(X_k) + \sum_{k \neq l} Cov(X_k, X_l)$$

Loi Bernoulli

On considere une variable aléatoire discrète X qui prend la valeur 1 avec la probabilité p, et 0 avec la probabilité $1-p$.

- $\Omega = \{0, 1\}$
- **Fonction de masse :** $P(X = x) = p^x(1-p)^{1-x}$, $x \in \{0, 1\}$
- **Espérance :** $\mathbb{E}[X] = 0(1-p) + 1p = p$
- **Variance :**
  $Var[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = [0(1-p) + 1p] - p^2 = p - p^2 = p(1-p)$

Exemples : toutes sortes d'expériences avec deux issues possibles (comme lancer d'une pièce)

Exercice

La probabilité que Toto passe un examen est égale à p. En total, il doit écrire n examens. Considerons la variable $X = \{\text{nombre d'examens passés}\}$.

- Écrire la fonction de masse de X pour $n = 2$. Trouver $\mathbb{E}[X]$ et $Var[X]$.
- Écrire la fonction de masse de X pour $n = 3$. Trouver $\mathbb{E}[X]$ et $Var[X]$.
- En regardant vos réponses, est-ce que vous pouvez définir la loi de X pour n'importe quel n ?
- Calculer la probabilité que Toto échoue $k-1$ examens, mais réussi le $k-$ème.
- Toto doit écrire une infinité des examens. Considerons une variable $Y = \{\text{première examen passé}\}$. Calculer $\mathbb{E}[Y]$ et $Var[Y]$.

Loi Binomiale

Considerons n épreuves identiques et indépendantes de Bernoulli. La variable $X \sim B(n, p)$ qui correspond au nombre de fois où un événement de probabilité p se produit au cours de cette suite d'épreuves suit **la loi Binomiale**.

- $\Omega = \{1, \ldots, n\}$
- **Fonction de masse :** $P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$
- **Espérance :** $\mathbb{E}[X] = np$
- **Variance :** $Var[X] = np(1 - p)$

Loi géométrique

Une variable aléatoire X de loi géométrique $\mathcal{G}(p)$ correspond au rang de la première occurence d'un événement de probabilité p dans une suite d'épreuves indentiques et indépendantes de Bernoulli.

- $\Omega = \mathbb{N}^+$
- **Fonction de masse :** $P(X = k) = (1 - p)^{k-1}p$
- **Espérance :** $\mathbb{E}[X] = \frac{1}{p}$
- **Variance :** $Var[X] = \frac{1-p}{p^2}$

Loi uniforme discrète

Une variable aléatoire X qui prend une valeur dans un ensemble fini de n valeurs possibles suit une loi uniforme lorsque la probabilité que X prenne n'importe quelle valeur dans cet ensemble est égale à $\frac{1}{n}$.

- $\Omega = \{1, \ldots, n\}$
- **Fonction de masse :** $P(X = k) = \frac{1}{n}$
- **Espérance :** $\mathbb{E}[X] = \frac{n+1}{2}$
- **Variance :** $Var[X] = \frac{n^2-1}{12}$

Loi de Poisson

Si le nombre moyen d'occurrences dans un intervalle de temps fixé est λ, alors la probabilité qu'il existe exactement k occurences est donné par la fonction de masse de la loi de Poisson.

- $\Omega = \{1, \ldots, n\}$
- **Fonction de masse :** $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$
- **Espérance :** $\mathbb{E}[X] = \lambda$
- **Variance :** $Var[X] = \lambda$

Comment simuler les variables discrètes dans R ?

- **Loi Bernoulli** : `rbinom(N,1,p)`
- **Loi Binomiale** : `rbinom(N,n,p)`
- **Loi géométrique** : `rgeom(N,n)`
- **Loi uniforme discrète** : `sample(1:n,N,replace = TRUE)`
- **Loi de Poisson** : `rpois(N,lambda)`,

où N — taille d'échantillon

Exercice

La probabilité que Toto passe un examen est égale à p_T. La probabilité que Clara passe un examen est égale à p_C. Ils ont du écrire n examens chaqun. Calculer la probabilité que Toto a passé plus d'examens que Clara.

Indications :

- Calculer la probabilité que Toto a passé **exactement** k examens. Idem pour Clara.
- Calculer la probabilité qu'il a passé **strictement moins** de k examens. Idem pour Clara.
- En utilisant la formule des probabilités totales, resoudre l'exercice.