

2. Variables aléatoires unidimensionnelles

MTH2302D

S. de Montigny en collaboration avec S. Le Digabel
École Polytechnique de Montréal

A2017

(v1)

Plan

1. Définitions
2. Variables aléatoires discrètes (masse)
3. Variables aléatoires continues (densité)
4. Fonction de répartition
5. Caractéristiques d'une distribution
6. Fonctions d'une variable aléatoire
7. Distribution d'une fonction d'une variable aléatoire
8. Espérance et variance de fonctions de v.a.

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Exemple 1

On lance une pièce trois fois et on note le résultat. L'espace échantillon de cette expérience aléatoire est

$$\Omega = \{PPP, PPF, PFP, FPP, PFF, FPF, FFP, FFF\}.$$

Soit X : le nombre de P (piles) obtenus.

X associe un nombre réel $X(s)$ à chaque résultat $s \in \Omega$.

L'ensemble des valeurs possibles pour X est $R_X = \{0, 1, 2, 3\}$.

On définit une probabilité sur X :

$$P(X \in B) = P(\{s \in \Omega \mid X(s) \in B\}) \text{ avec } B \subseteq R_X.$$

Définition

Une *variable aléatoire* (v.a.) X d'un espace échantillon Ω est une fonction qui associe à chaque résultat $s \in \Omega$ un nombre réel $x = X(s)$.

Définition

L'ensemble des valeurs possibles pour une v.a. X est appelé le *support* de X , dénoté R_X .

Définition

- Si le support d'une v.a. X est dénombrable (ou fini) alors X est *discrète*.
- Si le support d'une v.a. X est une collection d'intervalles alors X est *continue*.

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Fonction de masse

Soit X une v.a. discrète de support $R_X = \{x_1, x_2, x_3, \dots\}$.

La *fonction de masse* (de probabilité) de X est la fonction p_X définie par

$$p_X(x) = P(X = x) \text{ pour tout } x \in R_X .$$

La fonction de masse satisfait

- $0 \leq p_X(x) \leq 1$ pour tout $x \in R_X$.
- $\sum_{x_i \in R_X} p_X(x_i) = 1$.
- $P(X \in B) = \sum_{x_i \in B} p_X(x_i)$ (avec $B \subseteq R_X$).

Exemple 2

Une boîte contient 5 DVDs parmi lesquels 2 sont défectueux. Un échantillon de 2 disques est prélevé (sans remise) de la boîte.

Soit X : le nombre de DVDs défectueux dans l'échantillon.

1. Déterminer la fonction de masse de la v.a. X .
2. Évaluer les probabilités $P(X > 1)$, $P(X \geq 1)$, $P(X < 2)$.

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Fonction de densité

Soit X une v.a. continue de support R_X . La *fonction de densité* (de probabilité) de X est la fonction f_X définie par

- $f_X(x) \geq 0$ pour tout $x \in \mathbb{R}$.
- $P(X \in B) = \int_B f_X(x) dx$ pour tout $B \subset \mathbb{R}$.

En particulier

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

$$(= P(a < X \leq b) = P(a \leq X < b) = P(a < X < b)).$$

Fonction de densité (suite)

La fonction de densité satisfait

$$\int_{-\infty}^{\infty} f_X(x) dx = 1.$$

Différences avec la fonction de masse :

- On peut avoir $f_X(x) > 1$.
- $f_X(a)$ ne correspond pas à $P(X = a)$.
- $P(X = a) = \int_a^a f_X(x) dx = 0$. La probabilité que X prenne la valeur isolée a est nulle.
- On a plutôt $\varepsilon f_X(a) \simeq P(a - \varepsilon/2 \leq X \leq a + \varepsilon/2)$ qui correspond à la probabilité que X prenne sa valeur dans $[a - \frac{\varepsilon}{2}; a + \frac{\varepsilon}{2}]$.

Exemple 3

L'erreur commise lors de la mesure du diamètre d'une pièce produite en série est approximée par une v.a. X (en mm) dont la fonction de densité est

$$f(x) = \begin{cases} c(1 - x^2) & \text{si } -1 < x < 1, \\ 0 & \text{sinon.} \end{cases}$$

1. Déterminer la valeur du paramètre c .
2. Calculer la probabilité que l'erreur d'une mesure soit supérieure à $1/2$ en valeur absolue.

Exemple 4

Soit X la durée de vie d'une ampoule. La fonction de densité de la v.a. X est

$$f(t) = \begin{cases} \lambda e^{-\lambda t} & \text{si } t \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

où λ est le *taux de défaillance*.

1. Déterminer, en fonction de λ , la probabilité que l'ampoule fonctionne pendant au moins T heures.
2. Calculer la probabilité qu'une ampoule ayant fonctionné pendant 100 heures fonctionne encore 150 heures additionnelles.
3. Sachant que pour ce type d'ampoule $P(X \geq 100) = 0.99$, déterminer λ .

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Fonction de répartition

Soit X une variable aléatoire (discrète ou continue). La *fonction de répartition* F_X de X est définie par

$$F_X(x) = P(X \leq x) \quad \text{pour tout } x \in \mathbb{R}.$$

Propriétés de la fonction de répartition :

1. $0 \leq F_X(x) \leq 1$ pour tout $x \in \mathbb{R}$.
2. F_X est non décroissante.
3. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow +\infty} F_X(x) = 1$.

Une formule utile : si $a < b$ alors

$$P(a < X \leq b) = F_X(b) - F_X(a).$$

Si X est une v.a. discrète

Soit X une variable discrète, $R_X = \{x_1, x_2, x_3, \dots\}$ son support et p_X sa fonction de masse. Alors

- La fonction F_X est une fonction en escalier.
- $$F_X(x) = P(X \leq x) = \sum_{x_i \leq x} p_X(x_i).$$
- Si les éléments de R_X sont triés ($x_1 < x_2 < \dots$), alors
$$\begin{cases} p_X(x_1) = F_X(x_1) , \\ p_X(x_i) = F_X(x_i) - F_X(x_{i-1}) \text{ pour tout } x_i \in R_X \text{ avec } i \geq 2. \end{cases}$$

Exemple 5

Une boîte contient 5 DVDs parmi lesquels 2 sont défectueux. Un échantillon de 2 disques est prélevé (sans remise) de la boîte.

Soit X : le nombre de DVDs défectueux dans l'échantillon.

Déterminer la fonction de répartition de la v.a. discrète X .

Si X est une v.a. continue

Soit X une v.a. continue et f_X sa fonction de densité. Alors

$$F_X(x) = P(X \leq x) = \int_{-\infty}^x f_X(t) dt.$$

Par le théorème fondamental du calcul :

$$\frac{d}{dx} F_X(x) = f_X(x)$$

et F_X est une fonction continue.

Exemple 6

L'erreur commise lors de la mesure du diamètre d'une pièce produite en série est approximée par une v.a. X (en mm) dont la fonction de densité est

$$f(x) = \begin{cases} c(1 - x^2) & \text{si } -1 < x < 1, \\ 0 & \text{sinon.} \end{cases}$$

Déterminer la fonction de répartition de la v.a. continue X .

Notion de distribution

La **distribution** d'une variable aléatoire :

- est un terme général pour décrire le modèle suivi par la v.a.
- englobe les notions de densité/masse et de répartition.

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Espérance mathématique

Soit X une variable aléatoire. L'*espérance mathématique* (ou *moyenne*) de X est

- $\mu = \sum_{x_i \in R_X} x_i p_X(x_i)$ si X est discrète.
- $\mu = \int_{-\infty}^{+\infty} x f_X(x) dx$ si X est continue.

Autres notations : $\mu = E(X) = \mathbb{E}[X]$.

Exemple 7

Une boîte contient 5 DVDs parmi lesquels 2 sont défectueux. Un échantillon de 2 disques est prélevé (sans remise) de la boîte.

Soit X : le nombre de DVDs défectueux dans l'échantillon.

Calculer l'espérance de la v.a. discrète X .

Exemple 8

L'erreur commise lors de la mesure du diamètre d'une pièce produite en série est approximée par une v.a. X (en mm) dont la fonction de densité est

$$f(x) = \begin{cases} \frac{3}{4}(1 - x^2) & \text{si } -1 < x < 1, \\ 0 & \text{sinon.} \end{cases}$$

Calculer l'espérance de la v.a. continue X .

Variance et écart-type

Soit X une variable aléatoire. La *variance* de X est

- $\sigma^2 = \sum_{x_i \in R_X} (x_i - \mu)^2 p_X(x_i)$ si X est discrète.
- $\sigma^2 = \int_{-\infty}^{+\infty} (x - \mu)^2 f_X(x) dx$ si X est continue.

L'*écart-type* de X est $\sigma = \sqrt{\sigma^2}$.

Autres notations : $\sigma^2 = \text{Var}(X) = V(X) = \mathbb{V}(X)$.

Variance (suite)

- $$\begin{aligned} V(X) &= E[(X - E[X])^2] \\ &= E[X^2] - (E[X])^2. \end{aligned}$$

- Si X est discrète,

$$V(X) = \sum_{x_i \in R_X} (x_i - \mu)^2 p_X(x_i) = \left(\sum_{x_i \in R_X} x_i^2 p_X(x_i) \right) - \mu^2.$$

- Si X est continue,

$$V(X) = \int_{-\infty}^{+\infty} (x - \mu)^2 f_X(x) dx = \left(\int_{-\infty}^{+\infty} x^2 f_X(x) dx \right) - \mu^2.$$

Exemple 9

Montrer que dans l'exemple 7, $\sigma^2 = 0.36$, et que dans l'exemple 8, $\sigma^2 = 0.2$.

Inégalité de Bienaymé-Tchebychev

Théorème

Si $E(X) = \mu$ et $V(X) = \sigma^2$, alors, pour tout $a > 0$, on a

$$P(\mu - a\sigma \leq X \leq \mu + a\sigma) > 1 - 1/a^2.$$

Moments d'ordre supérieur

Soit X une variable aléatoire et $k \geq 0$ un entier.

1. Le k^e moment de X par rapport à l'origine est

$$- \mu'_k = \sum_{x_i \in R_X} x_i^k p_X(x_i) \text{ si } X \text{ est discrète.}$$

$$- \mu'_k = \int_{-\infty}^{+\infty} x^k f_X(x) dx \text{ si } X \text{ est continue.}$$

2. Le k^e moment de X par rapport à la moyenne μ est

$$- \mu_k = \sum_{x_i \in R_X} (x_i - \mu)^k p_X(x_i) \text{ si } X \text{ est discrète.}$$

$$- \mu_k = \int_{-\infty}^{+\infty} (x - \mu)^k f_X(x) dx \text{ si } X \text{ est continue.}$$

Moments d'ordre supérieur (suite)

3. Moyenne : $\mu = \mu'_1$.

4. Variance : $\sigma^2 = \mu_2$.

5. Relation entre μ_k et μ'_k :

$$\mu_k = \sum_{j=0}^k (-1)^j \binom{k}{j} \mu^j \mu'_{k-j} .$$

Exemple 10

Utiliser ces relations pour montrer que $\sigma^2 = \mu'_2 - \mu^2$.

Coefficients de forme

- **Coefficient d'asymétrie** (*skewness*) :
 - $\gamma_1 = \mu_3/\sigma^3$.
 - Si $\gamma_1 > 0$, la distribution est étalée vers la droite. Si $\gamma_1 < 0$, c'est vers la gauche. Si la distribution est symétrique, alors $\gamma_1 = 0$. L'inverse n'est pas vrai.
- **Coefficient d'aplatissement** (*kurtosis*) :
 - $\beta_2 = \mu_4/\sigma^4$

Médiane

- La médiane de la v.a. X est une valeur \tilde{x} telle que

$$P(X \leq \tilde{x}) \geq 1/2 \text{ et } P(X \geq \tilde{x}) \geq 1/2$$

ou bien

$$P(X \leq \tilde{x}) \geq 1/2 \text{ et } P(X < \tilde{x}) \leq 1/2.$$

- Utile si on a de grandes valeurs dans R_X car cette valeur est moins influencée que la moyenne par les valeurs extrêmes.
- On n'a pas forcément $\tilde{x} \in R_X$.
- Si X est discrète, alors \tilde{x} peut ne pas être unique.
- Si X est continue, \tilde{x} est unique et peut être trouvée avec $P(X \leq \tilde{x}) = P(X \geq \tilde{x}) = 1/2$.
- Dans le cas continu, la médiane est un cas particulier de la notion de *quantile*.

Les centiles (quantiles ou percentiles)

On considère une v.a. X **continue** et p un nombre réel avec $0 \leq p \leq 1$. On appelle le $100p$ -ième quantile (ou quantile d'ordre p , ou centile, ou percentile) de X le nombre x_p tel que $P(X \leq x_p) = p$ et $P(X \geq x_p) = 1 - p$.

- Le 50-ième centile est la médiane : $\tilde{x} = x_{0.5}$.
- Quartiles : $Q_1 = x_{0.25}$, $Q_2 = \tilde{x} = x_{0.5}$, $Q_3 = x_{0.75}$.
- Écart interquartile : $\text{IQR} = Q_3 - Q_1$. Mesure la dispersion de la moitié de la distribution car $P(Q_1 \leq X \leq Q_3) = 50\%$.
- Parfois, x_p est noté le $100(1 - p)$ -ième quantile.

Autres caractéristiques

- **Mode de X** : c'est la (ou les) valeur x^* telle que $p_X(x^*) \geq p_X(x)$ pour tout $x \in R_X$ (cas discret). Dans le cas continu, remplacer p_X par f_X .
- **Étendue de X** : c'est $\max\{R_X\} - \min\{R_X\}$. Peut être $+\infty$.

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Événements équivalents

Soit X une variable aléatoire et H une fonction. Alors $Y = H(X)$ est une autre variable aléatoire.

On calcule la probabilité d'un événement C de R_Y en considérant l'événement équivalent B de R_X :

$$B = \{x \in R_X | H(x) \in C\}.$$

La probabilité est donnée par

$$P(Y \in C) = P(X \in B).$$

Exemple 11

Le diamètre d'un fil est une v.a. de densité

$$f_X(x) = \begin{cases} 200 & \text{si } 1 \leq x \leq 1.005, \\ 0 & \text{sinon.} \end{cases}$$

Quelle est la probabilité que l'aire de la section de ce fil soit inférieure à $1.01\pi/4$?

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Cas 1 : X et Y sont discrètes

Si X est une v.a. discrète et $Y = H(X)$ est discrète, alors la fonction de masse de Y est donnée par

$$p_Y(y) = P(Y = y) = \sum_{\{x \in R_X \mid H(x) = y\}} P(X = x).$$

Exemple 12

Soit X une v.a. dont la fonction de masse est donnée par la tableau suivant :

x	-2	-1	0	1	2
$p_X(x)$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{4}$	$\frac{1}{4}$

Déterminer la fonction de masse de $Y = X^2$.

Cas 2 : X est continue et Y est discrète

Si X est une v.a. continue et $Y = H(X)$ est discrète alors la fonction de masse de Y est donnée par

$$p_Y(y) = P(Y = y) = \int_{\{x \in R_X | H(x) = y\}} f_X(x) dx.$$

Exemple 13

Pour 1 kg d'un certain alliage la teneur en magnésium est une v.a. X de densité

$$f_X(x) = \begin{cases} \frac{x}{8} & \text{si } 0 \leq x \leq 4, \\ 0 & \text{sinon.} \end{cases}$$

Chaque kg de cet alliage génère une perte de 100\$ si la teneur est inférieure à 2, aucun profit si la teneur est supérieure à 2 mais inférieure à 3, et un profit de 400\$ si la teneur est supérieure à 3.

On considère 1 kg de cet alliage.

Déterminer la fonction de masse du profit Y et le profit moyen $E(Y)$.

Cas 3 : X et Y sont continues

Si X est une v.a. continue et $Y = H(X)$ est continue alors il y a deux façons de déterminer la fonction de densité de Y , dépendamment de la fonction H .

Première façon, qui s'applique toujours.

1. Déterminer la fonction de répartition de Y :

$$F_Y(y) = P(Y \leq y) = P(H(X) \leq y).$$

2. Obtenir la densité f_y en dérivant F_Y :

$$f_Y(y) = \frac{d}{dy} F_Y(y).$$

Cas 3 : X et Y sont continues (suite)

Deuxième façon, si H est strictement croissante ou décroissante.

Dans ce cas, la densité de Y est donnée par

$$f_Y(y) = f_X(x) \left| \frac{dx}{dy} \right|$$

où $x = H^{-1}(y)$ et $f_X(x)$ est exprimée en fonction de y .

Cas 3 : X et Y sont continues (suite)

Exemple 14

Première façon.

Un courant électrique d'intensité aléatoire X traverse un appareil de résistance constante r . La puissance obtenue est alors donnée par la v.a. $Y = rX^2$ (loi d'Ohm).

Supposons que la fonction de densité de X soit

$$f_X(x) = \begin{cases} 2xe^{-x^2} & \text{si } x \geq 0, \\ 0 & \text{sinon.} \end{cases}$$

Déterminer la fonction de densité de Y .

Cas 3 : X et Y sont continues (suite)

Exemple 15

Deuxième façon.

Soit X une v.a. de densité

$$f_X(x) = \begin{cases} e^{-x} & \text{si } x > 0, \\ 0 & \text{sinon} \end{cases}$$

et $Y = 3X - 2$.

Déterminer la fonction de densité de Y .

1. Définitions

2. Variables aléatoires discrètes (masse)

3. Variables aléatoires continues (densité)

4. Fonction de répartition

5. Caractéristiques d'une distribution

6. Fonctions d'une variable aléatoire

7. Distribution d'une fonction d'une variable aléatoire

8. Espérance et variance de fonctions de v.a.

Définition

Soit X et $Y = H(X)$ des v.a. L'*espérance mathématique* de Y est

- $E[H(X)] = \sum_{x_i \in R_X} H(x_i)p_X(x_i)$ si X est discrète.
- $E[H(X)] = \int_{-\infty}^{\infty} H(x)f_X(x)dx$ si X est continue.

Cas particuliers

1. Si $H(x) = x$, alors $E[H(X)] = \mu$ (moyenne de X).
2. Si $H(x) = (x - \mu)^2$, alors $E[H(X)] = \sigma^2$ (variance de X).
3. Si $H(x) = (x - \mu)^k$, alors $E[H(X)] = \mu_k$ (moments par rapport à la moyenne).
4. Si $H(x) = x^k$, alors $E[H(X)] = \mu'_k$ (moments par rapport à l'origine).
5. Si $H(x) = 1_A(x) = \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases}$, alors $E[H(X)] = P(X \in A)$.

Variance

Soit X une variable aléatoire et $Y = H(X)$. La *variance* de Y est

$$\begin{aligned} V[H(X)] &= E \left[(H(X) - E[H(X)])^2 \right] \\ &= E \left[(H(X))^2 \right] - (E[H(X)])^2. \end{aligned}$$

Cas particulier important

Si $H(x) = ax + b$ (H est linéaire) alors

- $E(aX + b) = aE(X) + b$.
- $V(aX + b) = a^2V(X)$.

Exemple 16

Soit X une variable aléatoire de densité

$$f_X(x) = \begin{cases} \frac{1}{2} & \text{si } -1 \leq x \leq 1, \\ 0 & \text{sinon.} \end{cases}$$

Si $Y = X^2$, calculer $E[Y]$ et $V[Y]$.