

Module 2 - Variables aléatoires à une dimension

STT-1000 Probabilités et statistique

Département de mathématiques et de statistique



UNIVERSITÉ
LAVAL

Introduction

Exemple 1

Le résultat d'une expérience aléatoire est souvent un nombre réel, ou encore on assigne une valeur réelle aux résultats possibles d'une expérience.

$\mathcal{E}$: On tire une pièce de monnaie 2 fois
 $\mathcal{U} = \{FF, FP, PF, PP\}$.

Exemple de valeur associée à chaque résultat : le nombre de fois où Pile apparaît.
On peut noter cette variable X .

$X = 0$ si le résultat est FF,

$X = 1$ si le résultat est FP ou PF,

$X = 2$ si le résultat est PP.

Variable aléatoire

Soit $\mathcal{E}$ une expérience aléatoire ayant comme espace échantillonnal $\mathcal{U}$.

Variable aléatoire

Soit X , une fonction qui associe un nombre réel $X(e)$ à chaque élément $e \in \mathcal{U}$. Alors on appelle X une **variable aléatoire**.

L'ensemble des valeurs que peut prendre X est noté $\mathbb{R}_X$ et est appelé **image de X** .

On note habituellement une v.a. par une lettre majuscule et les valeurs prises par cette v.a. par une lettre minuscule. On écrira par exemple $P(X = x)$ pour dénoter la probabilité que la variable aléatoire X prenne la valeur x .

Exemple 2

QUIZ!

$\mathcal{E}$: on lance un dé régulier et on observe la face du dessus.

$$\mathcal{U} = \{1, 2, 3, 4, 5, 6\}.$$

On peut définir une infinité de variables aléatoires.

Voici 2 exemples.

- X est la variable aléatoire qui donne le résultat observé.

$$X(e) = e, \forall e \in \mathcal{U}$$

$$\mathbb{R}_X = \{1, 2, 3, 4, 5, 6\}$$

- Je gagne 1 \$ si je roule 4, 5 ou 6.

Je perds 1 \$ si je roule 1, 2 ou 3.

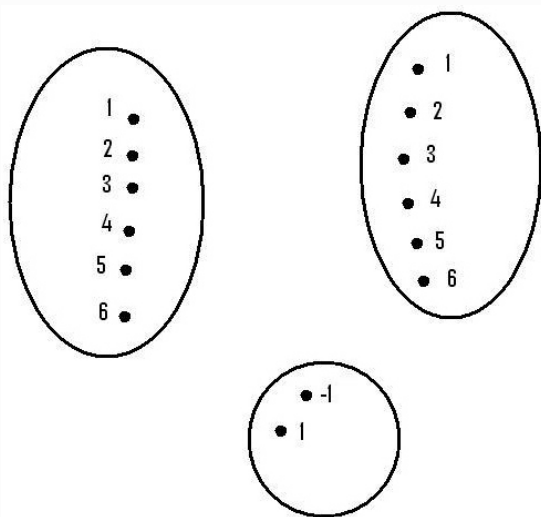
Y est la variable aléatoire qui représente mon gain.

$$Y(e) = \begin{cases} -1 & \text{si } e \in \{1, 2, 3\} \\ 1 & \text{si } e \in \{4, 5, 6\} \end{cases}$$

$$\mathbb{R}_Y = \{-1, 1\}.$$

Exemple 2 (suite)

e	$X(e)$	$Y(e)$
1		
2		
3		
4		
5		
6		



Exemple 2 : probabilités

Quelle est la probabilité que je perde 1 \$ en jouant ?

e	$X(e)$	$Y(e)$	$P(e)$
1			1/6
2			1/6
3			1/6
4			1/6
5			1/6
6			1/6

Fonction de répartition

Fonction de répartition

On peut spécifier complètement le comportement aléatoire de **toute** variable aléatoire X à l'aide d'une seule fonction.

Définition. Soit X une variable aléatoire. On appelle **fonction de répartition** (CDF, de *cumulative distribution function*) la fonction

$$F_X : \mathbb{R} \rightarrow [0, 1]$$

qui associe la valeur

$$F_X(x) = P(X \leq x),$$

à tout $x \in \mathbb{R}$.

Retour à l'exemple 2

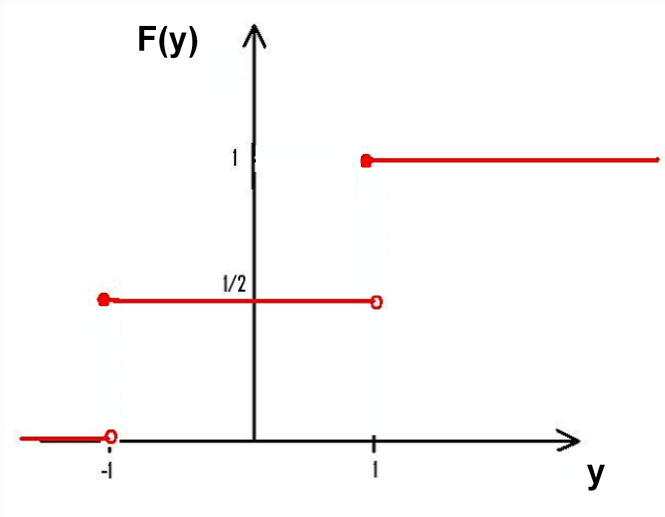
Quelle est la fonction de répartition du gain Y lors du lancer d'un dé équilibré?

On sait que

$$P(Y = -1) = 1/2$$

$$P(Y = 1) = 1/2$$

Retour à l'exemple 2 (suite)



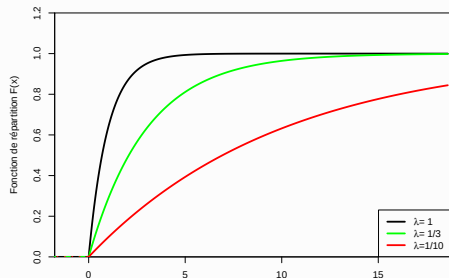
Exemple 3

QUIZ!

Soit X la quantité de pluie (en mm) reçue à une station météo dans un mois donné. Un modèle possible pour le comportement aléatoire de X est

$$F(x) = P(X \leq x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-\lambda x} & \text{si } x \geq 0, \end{cases}$$

où λ est une constante positive, que l'on suppose connue.



Exemple 3 (suite)

Supposons que $\lambda = 1/10$. Calculez ces probabilités.

- $P(X \leq 15)$
- $P(X > 10)$
- $P(5 < X \leq 10)$

Retour à l'exemple du dé (Exemple 2)

QUIZ!

Quelle est la probabilité d'obtenir un gain supérieur à 0,75 \$?

Répondez à l'aide de la fonction de répartition.

Propriétés d'une fonction de répartition

1. $0 \leq F_X(x) \leq 1$ pour tout $x \in \mathbb{R}$
 2. $\lim_{x \rightarrow -\infty} F_X(x) = 0$ et $\lim_{x \rightarrow \infty} F_X(x) = 1$
 3. $F_X(x)$ est **non décroissante** (si $x_1 \leq x_2$, alors $F_X(x_1) \leq F_X(x_2)$)
 4. $F_X(x)$ est **continue à droite**, $\left(\lim_{s \rightarrow 0+} F_X(x + s) = F_X(x), \forall x \in \mathbb{R} \right)$
- Pour vérifier si une fonction est une fonction de répartition, on vérifie si elle remplit les 4 propriétés ci-dessus.
 - De la définition d'une fonction de répartition, on a que pour tout $a < b$, $P(a < X \leq b) = F_X(b) - F_X(a)$.

Variables aléatoires discrètes et continues

- Si $F_X(x)$ est une fonction en escalier continue à droite, alors on dit que X est une variable aléatoire **discrète**. L'ensemble des valeurs possibles est **dénombrable**.
- Si $F_X(x)$ est une fonction continue, alors on dit que X est une variable aléatoire **continue**. L'ensemble des valeurs possibles est **indénombrable**.

Variables aléatoires discrètes

Fonction de probabilité/masse

Définition. Soit X une v.a. **discrète**. On appelle **fonction de probabilité** (ou **fonction de masse**, ou **fonction de masse de probabilité**; PMF, de *probability mass function*) de X la fonction

$$p_X : \mathbb{R} \rightarrow [0, 1]$$

qui associe la valeur

$$p_X(x) = P(X = x)$$

à tout $x \in \mathbb{R}$.

Exemples

Retournons à l'exemple du dé.
Trouvons p_X et p_Y .

Exemples - suite

Propriétés de la fonction de probabilité

Une fonction p_X peut être une fonction de probabilité pour une v.a. discrète X **si et seulement si** elle satisfait deux propriétés :

- 1.

- 2.

Variables aléatoires continues

Fonction de densité

Définition. Soit X une v.a. continue. On appelle **fonction de densité** (PDF, de *probability density function*) de X la fonction qui associe la valeur

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

à tout $x \in \mathbb{R}$.

Lien entre la fonction de répartition $F_X(x)$ et la fonction de densité $f_X(x)$

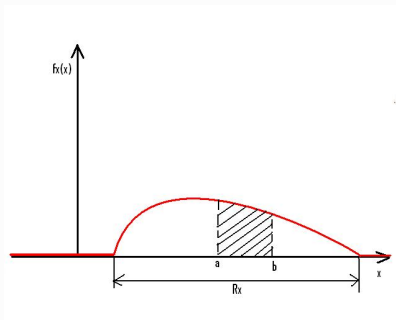
Par la définition de la fonction de densité $f_X(x)$, on remarque :

$$F_X(x) = \int_{-\infty}^x f_X(u) \, du$$

$$P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(u) \, du.$$

On en déduit donc que la probabilité d'une variable aléatoire continue peut se calculer avec l'intégrale de $f_X(x)$.

Probabilité = aire sous la densité



$f_X(x)$ n'est **pas** une probabilité, mais son intégrale sur une région donne la probabilité que X prenne une valeur dans cette région.

Ceci veut dire que $f_X(x) \neq P(X = x)$.

Propriétés d'une fonction de densité

1. $f_X(x) \geq 0$ pour tout $x \in \mathbb{R}$
2. $\int_{-\infty}^{\infty} f_X(x) dx = 1$
3. $f_X(x)$ est continue par morceaux
4. si $x \notin \mathbb{R}_X$, alors $f_X(x) = 0$

Comme $\int_a^a f_X(u) du = 0$, alors $P(X = a) = 0$ pour tout $a \in \mathbb{R}$ dans le cas d'une v.a. continue. Ceci implique que

$$\left. \begin{array}{l} P(a < X < b) \\ P(a < X \leq b) \\ P(a \leq X < b) \\ P(a \leq X \leq b) \end{array} \right\} = \int_a^b f_X(u) du.$$

Exemple 4

QUIZ!

Soit une v.a. X continue pouvant prendre toute valeur dans l'intervalle $[0, 20]$. On vous dit que

- tous les intervalles de même longueur dans $[0, 10)$ sont équiprobables,
- tous les intervalles de même longueur dans $[10, 20]$ sont équiprobables,
- tous les intervalles dans $[10, 20]$ sont deux fois plus probables qu'un intervalle de même longueur dans $[0, 10)$.

Déterminez la densité de X , sa fonction de répartition, et évaluez $P(5 < X < 12)$.

Exemple 5 : Retour à l'exemple de la pluie

La fonction de répartition de la quantité de pluie mensuelle reçue à la station est définie par $F(x) = 1 - e^{-\lambda x}$ si $x \geq 0$, et 0 ailleurs.

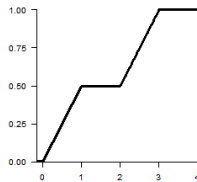
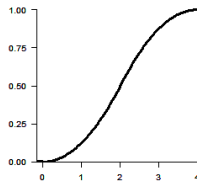
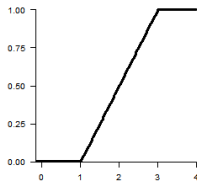
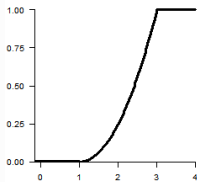
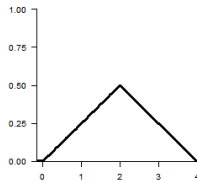
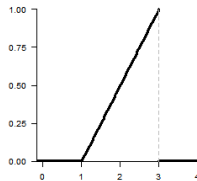
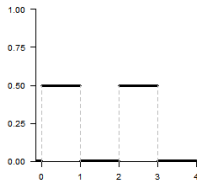
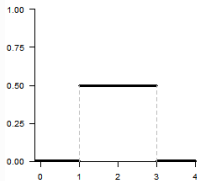
- Quelle est la densité de X ?
- Quelle est la probabilité que la quantité de pluie reçue excède 5 mm?

QUIZ!

- Quelle est la probabilité de recevoir entre 10 et 15 mm sachant que l'on recevra au moins 5 mm?

QUIZ!

Associez chaque densité avec la bonne répartition.



Espérance et variance d'une variable aléatoire

Aspects importants d'une distribution

Même si les fonctions de répartition, de densité ou de probabilité spécifient complètement la distribution d'une variable aléatoire, il est utile de s'intéresser aux caractéristiques qui résument les aspects principaux de son comportement :

- tendance centrale;
- dispersion;
- forme.

Notez qu'à partir de maintenant, quand il n'y aura qu'une seule variable aléatoire à l'étude, nous écrirons $F(x)$, $p(x)$ et $f(x)$ au lieu de $F_X(x)$, $p_X(x)$ et $f_X(x)$.

Mesure de la tendance centrale

Définition. L'**espérance** (moyenne théorique) d'une variable aléatoire X est donnée par

$$E(X) = \sum_{x \in \mathbb{R}_X} x p(x)$$

pour une variable aléatoire discrète,

ou

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx.$$

pour une variable aléatoire continue.

La notation la plus fréquente est $E(X) = \mu$.

Interprétation de l'espérance de X

C'est la valeur moyenne que l'on obtiendrait si on réalisait l'expérience aléatoire une infinité de fois.

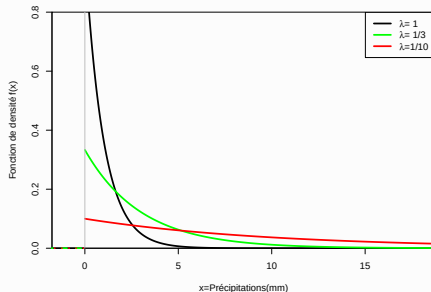
Pour une variable aléatoire continue, en pratique, nous n'observerons jamais une moyenne sur un nombre fini d'expériences aléatoires exactement égale à l'espérance.

Ceci n'est pas tout à fait vrai pour une variable aléatoire discrète. Un exemple ?

Exemple 6 : Retour à l'exemple de la pluie

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{si } x \leq 0 \end{cases}$$

Quelle quantité de pluie tombe-t-il en moyenne durant ce mois? (Répondez pour λ en général, sans fixer de valeur).



Variabilité d'une variable aléatoire

Deux variables aléatoires peuvent avoir la même espérance, mais l'une peut être beaucoup plus variable que l'autre, i.e., avoir tendance à prendre des valeurs plus éloignées de son espérance que l'autre.

Expérience : Rouler 30 dés réguliers.

X_1 = la valeur indiquée par le premier dé

X_2 = la moyenne des 30 dés (la somme divisée par 30).

On peut montrer que X_1 et X_2 ont la même espérance (3, 5).

Cependant, X_2 aura tendance à prendre des valeurs très proches de 3, 5, alors que X_1 peut prendre n'importe quelle valeur entre 1 et 6 avec la même probabilité.

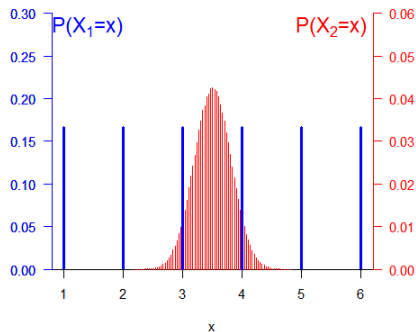
Petite simulation

Voici un exemple où on a simulé 12 fois l'expérience à l'aide de R :

	X1	X2
[1,]	2	3.633333
[2,]	6	3.666667
[3,]	4	3.933333
[4,]	4	3.633333
[5,]	5	3.833333
[6,]	6	3.600000
[7,]	6	3.400000
[8,]	3	3.166667
[9,]	2	3.766667
[10,]	5	3.066667
[11,]	3	3.566667
[12,]	1	3.300000

Fonctions de probabilité de X_1 et X_2

La figure ci-dessous montre la fonction de probabilité de X_1 en bleu et celle de X_2 en rouge. Même si les deux v.a. ont la même moyenne, on voit bien que la distribution de X_2 est beaucoup plus concentrée autour de 3,5 que celle de X_1 .



Mesure de dispersion d'une distribution

Définition. La **variance** d'une variable aléatoire X est donnée par

$$V(X) = \sum_{x \in \mathbb{R}_X} (x - \mu)^2 p(x)$$

pour une variable aléatoire discrète et

$$V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx.$$

pour une variable aléatoire continue.

La notation la plus fréquente est $V(X) = \sigma^2$.

Une quantité très utile qui découle de la variance est l'écart-type,

$$\sigma = \sqrt{V(X)} = \sqrt{\sigma^2}.$$

Exemple 7 : Retour à l'exemple de la pluie

La quantité de pluie reçue en un mois à cette station est-elle très variable ?

Espérance d'une fonction d'une variable aléatoire

Espérance d'une fonction d'une variable aléatoire

Espérance d'une fonction d'une variable aléatoire

Soit X une v.a. et posons $Y = H(X)$ une fonction de X . Alors l'**espérance de $H(X)$** (ou de Y) est définie par

$$E[Y] = E[H(X)] = \sum_{x \in \mathbb{R}_X} H(x)p(x)$$

ou

$$E[Y] = E[H(X)] = \int_{-\infty}^{\infty} H(x)f(x) dx.$$

Un cas particulier...

- $H(X) = (X - \mu)^2 \Rightarrow E[(X - \mu)^2] = V[X] = \sigma^2 = \text{variance de } X.$

Résultats utiles

Supposons que a et b sont des constantes.

- $E[aX + b] = aE[X] + b$

- $V[X] = E[X^2] - (E[X])^2$

- $V[aX + b] = a^2V[X]$

Prouvez $E[aX + b] = aE[X] + b$ dans le cas d'une variable aléatoire continue.

Prouvez $V[X] = E[X^2] - (E[X])^2$.

Prouvez $V[aX + b] = a^2V[X]$.

Exemple 8 : Retour à l'exemple de la pluie

Supposons que les revenus (en centaines de milliers de dollars) générés par un barrage sont bien modélisés par l'équation $R = 10 + 3X$ (où X est la quantité de pluie). Quelle est la valeur moyenne et l'écart-type de ces revenus ?

Rappel : $E(X) = 1/\lambda$ $V(X) = 1/\lambda^2$

Réponses des exemples

Réponses des exemples (1/2)

- **Exemple 2 (suite) (table $X(e)$, $Y(e)$ pour le dé) :** $X(e) = e$ pour $e = 1, \dots, 6$; $Y(e) = -1$ si $e \in \{1, 2, 3\}$, $Y(e) = 1$ si $e \in \{4, 5, 6\}$.
- **Exemple 2 : probabilités (perdre 1 \$) :** $P(Y = -1) = P(e \in \{1, 2, 3\}) = 3/6 = 1/2$.
- **Retour à l'exemple 2 (fonction de répartition de Y) :** $F_Y(y) = 0$ si $y < -1$; $F_Y(y) = 1/2$ si $-1 \leq y < 1$; $F_Y(y) = 1$ si $y \geq 1$.
- **Exemple 3 (suite) ($\lambda = 1/10$) :** $P(X \leq 15) = 1 - e^{-1,5} \approx 0,777$; $P(X > 10) = e^{-1} \approx 0,368$; $P(5 < X \leq 10) = e^{-0,5} - e^{-1} \approx 0,239$.
- **Exemples (p_X , p_Y pour le dé) :** $p_X(1) = \dots = p_X(6) = 1/6$; $p_Y(-1) = p_Y(1) = 1/2$.
- **Propriétés de la fonction de probabilité :** (1) $p_X(x) \geq 0$ pour tout x ; (2) $\sum_x p_X(x) = 1$.

Réponses des exemples (2/2)

- **Exemple 5 (retour à la pluie) :** $f(x) = \lambda e^{-\lambda x}$ si $x > 0$ (**0 sinon**); $P(X > 5) = e^{-5\lambda}$.
- **Exemple 6 :** $E(X) = 1/\lambda$.
- **Exemple 7 :** $V(X) = 1/\lambda^2$; **plus λ est petit, plus X est variable.**
- **Preuves (« Résultats utiles ») :** découlent directement des définitions de $E(\cdot)$ et $V(\cdot)$ en substituant $aX + b$, puis, pour la deuxième, en développant $(X - \mu)^2 = X^2 - 2\mu X + \mu^2$.
- **Exemple 8 :** $E(R) = 10 + 3/\lambda$, $V(R) = 9/\lambda^2$, **donc l'écart-type est $3/\lambda$.**

Réponses des quiz

Réponses des quiz

- **Quiz (retour au dé, gain $> 0,75\$$) :** $P(Y > 0,75) = P(Y = 1) = 1/2$ (car $F_Y(0,75) = 1/2$).
- **Quiz (Exemple 4, densité sur $[0, 20]$) :** $f(x) = 1/30$ si $x \in [0, 10)$, $f(x) = 1/15$ si $x \in [10, 20]$, **0** sinon ; $F(x) = x/30$ si $0 \leq x < 10$, $F(x) = 1/3 + (x - 10)/15$ si $10 \leq x \leq 20$; $P(5 < X < 12) = 3/10 = 0,3$.
- **Quiz (pluie, $P(10 < X < 15 \mid X \geq 5)$) :** par la propriété d'absence de mémoire, $= e^{-5\lambda} - e^{-10\lambda}$.
- **Quiz (association densité-répartition) :** densité 1 (uniforme sur $[1, 3]$) $\leftrightarrow$ répartition 2 (rampe linéaire) ; densité 2 (deux blocs sur $[0, 1]$ et $[2, 3]$) $\leftrightarrow$ répartition 4 ; densité 3 (rampe croissante sur $[1, 3]$) $\leftrightarrow$ répartition 1 (courbe convexe) ; densité 4 (triangulaire sur $[0, 4]$) $\leftrightarrow$ répartition 3 (courbe en S).

Bibliographie

Ce document est rédigé avec Beamer ainsi qu'une classe .cls fournie par Jérôme Soucy. Les diapositives sont adaptées des diapositives créées par Anne-Sophie Charest pour le cours STT-1000 ainsi que celles créées par Thierry Duchesne et Emmanuelle Reny-Nolin pour le cours STT-1900.

Pour signaler une erreur dans ce document, veuillez écrire à l'adresse : julien.miron@mat.ulaval.ca.