Réseaux IP

110. Rappel de Probabilités: Variables aléatoires, Espérance et Variance

(Résumé destiné au cours de Réseaux IP)

Table des matières

1. Introduction

L'étude des performances des réseaux de téléinformatique et de télécommunication est largement basée sur le calcul des probabilités. Le but de ce document est de revoir les notions de base: variable aléatoire, espérance mathématique et variance qui sont indispensables à une bonne compréhension de la théorie du télétrafic.
On donne ici une approche non théorique. Une approche plus rigoureuse peut être trouvée dans de nombreux ouvrages sur les probabilités. Une brève liste de références est donnée à la fin de ce chapitre.
Pour visualiser ce qui se passe, des exemples sont donnés. De façon "classique", les exemples avec un ensemble de possibilités restreint se basent souvent sur les dés à jouer.

2. Probabilités et phénomènes aléatoires

Il est difficile de définir exactement ce qu'est une "probabilité".

Le calcul des **probabilités** est l'étude des phénomènes **non déterministes** aussi appelés phénomènes **aléatoires**. Ce sont les phénomènes (résultats d'expériences) dont le résultat n'est pas connu à priori.

Seule une probabilité peut être associée aux divers résultats possibles. La probabilité 0 est associée aux événements impossibles. La probabilité 1 est associée à l'événement certain. Pour chaque phénomène ou expérience, le **calcul des probabilités** attribue des "valeurs limites de fréquence relative" aux événements qu'on appelle **probabilité de ces événements**. On note cette probabilité Prob(événement) ou P(événement). Elle est comprise entre 0 et 1 (inclus).

Exemple
Si on jette une pièce de monnaie, il n'est pas certain que le côté pile va apparaître. Cependant, si la pièce est équilibrée, on peut dire que sur une longue série d'essais, la moitié environ des essais vont donner le côté pile. Il y a donc 50% de chances d'avoir le côté pile.

Deux événements qui ont la même probabilité sont dits **équiprobables**.

Une branche voisine des probabilités, la **statistique**, étudie l'estimation et le test de ces probabilités à partir de l'expérience. En simplifié, on peut dire que les probabilités étudient ce qui va se passer à partir des probabilités des événements élémentaires, alors que la statistique cherche à estimer ces probabilités à partir des résultats d'une expérience.

3. L'ensemble fondamental des événements

L'ensemble S de tous les résultats possibles S_i d'une expérience est appelé **ensemble fondamental**.

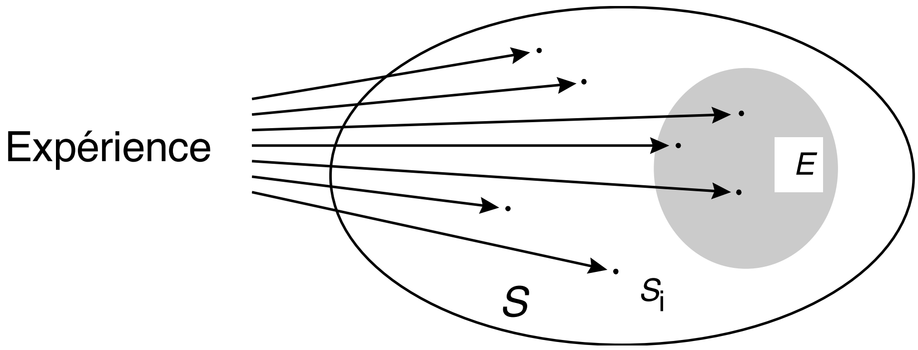


Fig. 3.1. Ensemble fondamental S, résultats S_i et événement E.

On peut définir des sous-ensembles de S. Par exemple, ci-dessus, E est un événement composé de trois résultats. L'ensemble vide f et l'ensemble S sont aussi des événements.
A chaque résultat (ou événement élémentaire), on peut associer une probabilité:

$$0 \leq \mathrm{Prob}(S_i) \leq 1 \tag{3.1}$$

La somme des probabilités de tous les résultats possibles est égale à 1:

$$\sum_i \mathrm{Prob}(S_i) = \mathrm{Prob}(S) = 1 \tag{3.2}$$

Exemple
Si on jette un dé, l'ensemble S est composé des résultats S = {1, 2, 3, 4, 5, 6}. L'événement "nombre pair" est composé de {2, 4, 6} et l'événement "nombre premier" de {1, 2, 3, 5}.

On peut combiner des événements pour former de nouveaux événements en suivant la théorie des ensembles. Notons que l'ensemble fondamental peut être fini ou infini. De même, il peut être discret ou continu.

Exemple
A È B est l'événement qui se produit si l'événement A ou l'événement B est réalisé. A ∩ B est l'événement qui se produit si A et B sont réalisés.

L'axiome de Bayes, nous donne la probabilité conditionnelle d'un événement *B* si l'événement *A* s'est produit:

$$\text{Prob}\,(B \textit{ sachant } A \textit{ réalisé}) = \text{Prob}(B|A) = \frac{\text{Prob}\,(A \cap B)}{\text{Prob}(A)} \quad ; \quad P(A) \neq 0 \qquad (3.2)$$

$$\text{Prob}(B|A) = \frac{\text{Prob}\,(B)}{\text{Prob}(A)} \quad si \quad B \subset A$$

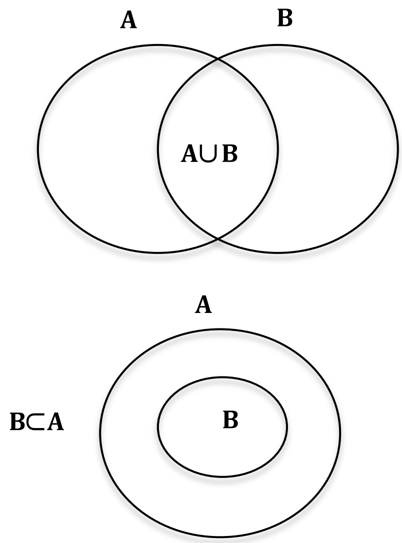


Fig. 3.2. *Axiome de Bayes*.

4. Variables aléatoires

Très souvent, lors d'un phénomène ou d'une expérience, on est intéressé à un nombre issu des résultats plutôt qu'au résultat lui-même.

Exemples

Lors d'un jeu de dés, on est souvent intéressé à la somme de points, pas au résultat de chaque dé.

Un pédiatre veut connaître le poids (en grammes) et la taille (en cm) des nouveaux-nés.

Dans un réseau téléinformatique, on est intéressé à connaître le temps de transfert moyen.

C'est là qu'interviennent les variables aléatoires: une **variable aléatoire** est une fonction qui associe un **nombre** à un **résultat** d'expérience.

C'est malheureusement une notion qui est souvent mal comprise et qui est entourée d'un flou pas toujours très artistique. Une des raisons pour cela est qu'une variable aléatoire n'est pas une variable mais une **fonction**. Une variable aléatoire est parfois abrégée V.A.

Définition

Soit S l'ensemble fondamental d'une expérience. Une **variable aléatoire** X attribue à chaque résultat de S un **nombre réel.**

Par convention, pour une variable aléatoire X (X majuscule), on note ces nombres réels par x (x minuscule) dans le cas continu et $x_1, x_2, x_3, \ldots, x_i, \ldots x_n$ dans le cas discret fini.

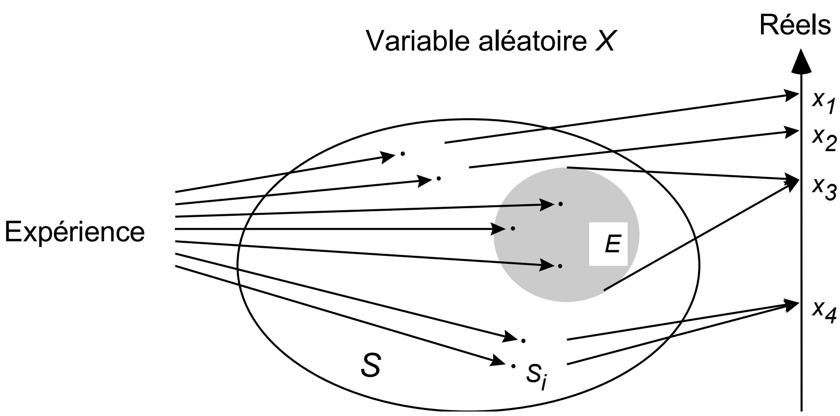


Fig. 4.1. Variable aléatoire X.

Exemple

On jette deux dés. L'ensemble fondamental S est composé des 36 résultats (1,1), (1,2), (1,3), ... (6,5) et (6,6). Ces résultats sont équiprobables. On définit une variable aléatoire Z qui attribue le produit z des deux chiffres obtenus à chaque résultat.

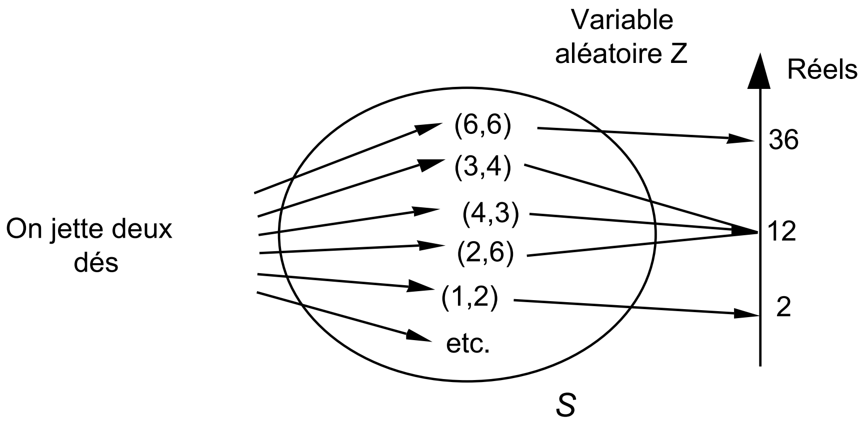


Fig. 4.2. Exemple: variable aléatoire Z.

Les variables aléatoires peuvent être discrètes (nombre de coups de téléphones dans une journée), continues (intervalles entre deux coups de téléphone), finies (nombre de paquets dans la mémoire d'un routeur) ou infinies (intervalle de temps entre deux paquets de la même source).

5. Probabilité d'une variable aléatoire

A chaque valeur (intervalle) de la **variable aléatoire** *X* sur l'axe réel, on peut attribuer une **probabilité** qui est celle de l'événement qui génère cette valeur. On la note

$$f\,(\text{valeur}) = \text{Prob}\,(X = \text{valeur}) \tag{5.1}$$

Dans le cas d'un ensemble fondamental continu, un intervalle des nombres réels correspond à un événement (qui peut être vide).

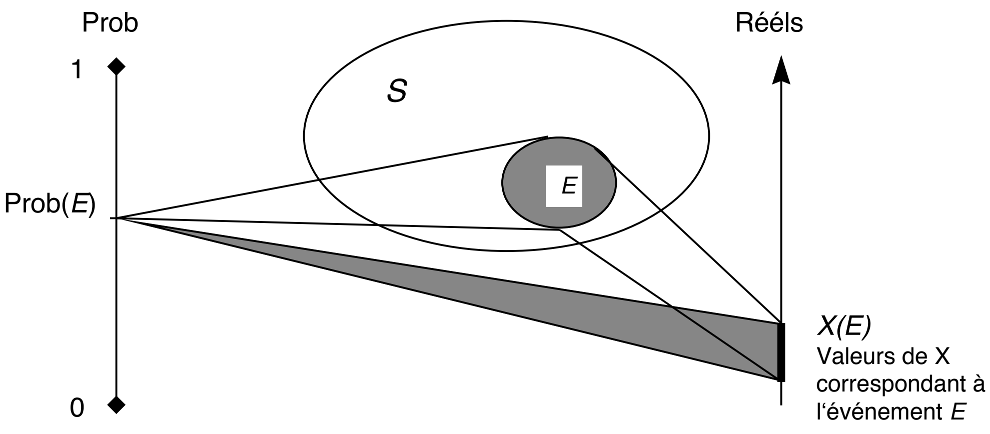


Fig. 5.1. Ensemble fondamental *S*, événement *E*, V.A. *X* et probabilités.

Exemple

On jette deux dés. On définit une variable aléatoire Y qui fait correspondre à chaque résultat (S est fini et discret) le maximum des deux chiffres:
(a,b) → max (a,b)
par exemple (3, 4) → 4
On obtient les probabilités associées en considérant les événements
f(1) = Prob (Y=1) = Prob (1,1) = 1/36
f(2) = Prob (Y=2) = Prob {(1,2), (2,1), (2,2)} = 3/36
f(3) = Prob (Y=3) = Prob {(1,3), (2,3), (3,3), (3,2), (3,1)} = 5/36
de même f(4) = 7/36, f(5) 9/36 et f(6) = 11/36

6. Fonction de répartition

La fonction de répartition *F(x)* d'une variable aléatoire *X* est la fonction

$$F(x) = \text{Prob}(X \leq x) \tag{6.1}$$

Cette fonction est aussi appelée fonction cumulative ou fonction de distribution. C'est la probabilité que la variable aléatoire *X* soit plus petite qu'une valeur de seuil *x*. On voit toute suite que $F(\infty) = 1$ et que $F(-\infty) = 0$. *F(x)* est donc une probabilité cumulative.

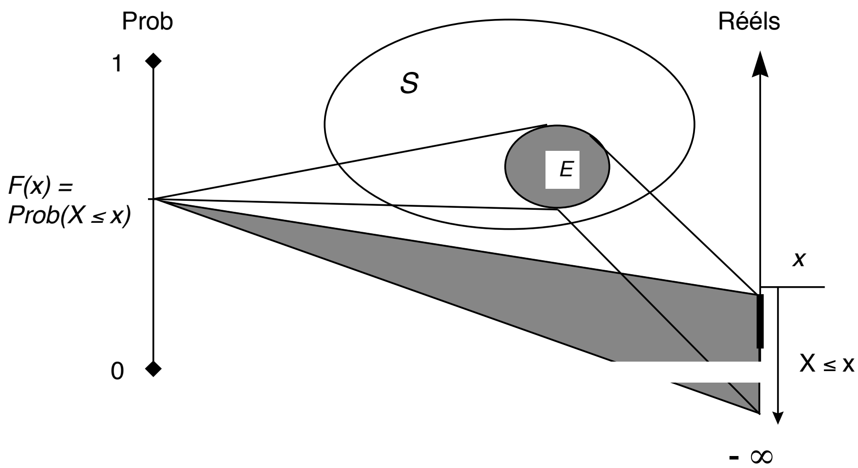


Fig. 6.1. Variable aléatoire X et Prob(X ≤ x).

Exemple

On jette un dé et on définit la variable aléatoire X comme $X(i) = x_i = 10i$ *où i est le chiffre obtenu. On attribue donc à chaque résultat un nombre qui est égal à 10 fois le chiffre obtenu. Notons que la variable aléatoire est discrète et finie.*
$X(1) = x_1 = 10$
$X(2) = x_2 = 20$, *etc.*
Prob(X=10) = Prob (1) = 1/6
Prob(X=20) = Prob (2) = 1/6, etc.

On a par exemple F(8) = 0, F(15) = 1/6, F(100) = 1

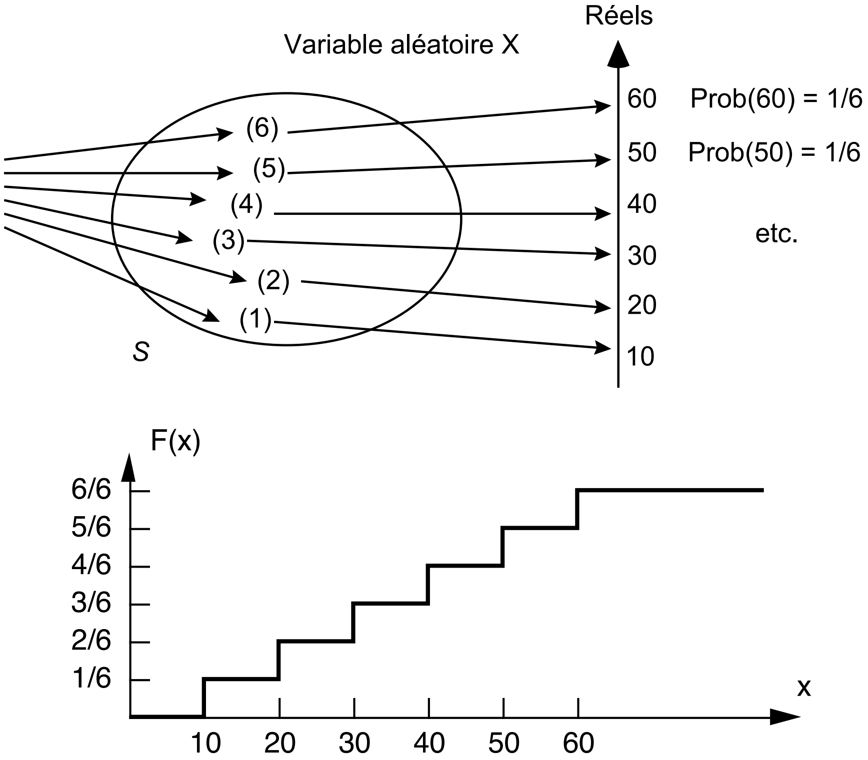


Fig. 6.2. Exemple de variable aléatoire et fonction de répartition associée.

7. Densité de probabilité

La **densité de probabilité** *f(x)* d'une variable aléatoire *X* indique l' "intensité de la probabilité" au voisinage de *x*. Elle est définie comme la dérivée de la fonction de répartition:

$$f(x) = \frac{dF(x)}{dx} \tag{7.1}$$

Cette définition s'applique aussi bien au cas discret qu'au cas continu. Dans le cas discret, on obtient des impulsions de Dirac aux valeurs discrètes de la variable aléatoire. On a immédiatement la réciproque

$$F(x) = \int_{-\infty}^{x} f(t)\,dt \quad ; \forall x \in Réels \tag{7.2}$$

Exemple

Pour l'exemple ci-dessus, la densité de probabilité peut être représentée de la façon suivante:

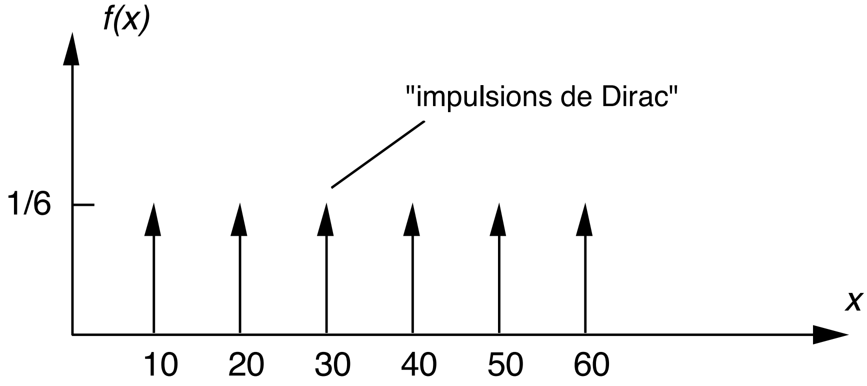


Fig. 7.1. Densité de probabilité f(x) correspondant à l'exemple de la Section 6.

Exemple

Dans le cas d'une variable aléatoire continue, on a par exemple la situation suivante, qui se rencontre souvent en pratique

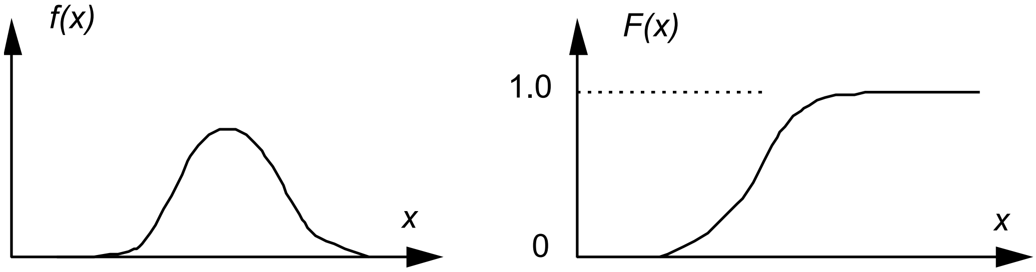


Fig. 7.2. Exemple de variable aléatoire continue: représentation de *f(x)* et *F(x).*

8. Espérance

L'espérance mathématique *E(X)* d'une variable aléatoire discrète finie *X* est la **moyenne pondérée** de la variable aléatoire en tenant compte de la probabilité de chaque valeur:

$$E(X) = x_1 f(x_1) + x_2 f(x_2) + \ldots = \sum_{i=1}^{n} x_i f(x_i) \tag{8.1}$$

Très souvent on utilise la notation $\mu = E(X)$. Parfois aussi $\overline{X} = E(X)$. L'espérance est **un nombre** (obtenu pour une fonction et une expérience), ce **n'est pas une fonction**. Dans le cas d'une variable aléatoire discrète infinie, on peut directement généraliser

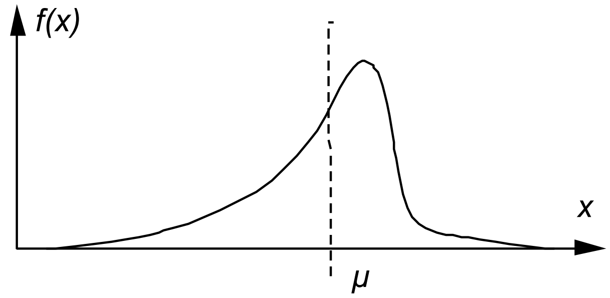
$$E(X) = \sum_{i=1}^{\infty} x_i f(x_i) \tag{8.2}$$

Exemple
Si la variable aléatoire discrète Z prend la valeur 2 avec une probabilité 0.9 et 3 avec une probabilité 0.1, l'espérance est μ = 2 ´ 0.9 + 3 ´ 0.1 = 2.1

Dans le cas d'une variable aléatoire continue il faut intégrer sur tout l'axe réel:

$$E(X) = \mu = \int_{-\infty}^{\infty} x\, f(x)\, dx \tag{8.3}$$

Exemple



Exemple
On donne la densité de probabilité suivante:

$$\begin{aligned} f(x) &= 0 \quad ; \quad x < 0 \\ f(x) &= e^{-x} \quad ; \quad x \geq 0 \end{aligned} \tag{8.4}$$

L'espérance est donnée par l'intégrale (8.3) qui peut être calculée par parties:

$$E(X) = \int_0^{\infty} x e^{-x} dx = -x e^{-x} \Big|_0^{\infty} + \int_0^{\infty} e^{-x} dx = 0 - e^{-x} \Big|_0^{\infty} = 1 \tag{8.5}$$

L'espérance correspond donc à la moyenne des valeurs de la variable aléatoire obtenue **si on répète inlassablement l'expérience**. Les événements rares, même si leur valeur s'écarte beaucoup de la moyenne, on relativement de la peine à influencer les événements fréquents.

On voit toute suite que dans le cas d'une variable aléatoire constante *X = Cte, E(X) = Cte*.

Exemple
On jette une paire de dés. L'espace S des événements (équiprobables) est

S = {(1,1), (1,2), (1,3), ...(6,6)}

On définit une V.A. discrète et finie Y qui fait correspondre à chaque événement (i, j) le maximum de (i, j). Sur l'axe réel, Y prend les valeurs {1, 2, 3, 4, 5, 6} avec les probabilités (voir Section 5)

f(1) = Prob (Y=1) = 1/36
f(2) = Prob (Y=2) = 3/36
f(3) = Prob (Y=3) = 5/36
f(4) = 7/36, f(5) = 9/36 et f(6) = 11/36

Alors, en utilisant la définition (8.1),

$$E(Y) = \sum_{i=1}^{6} y_i f(y_i)$$
$$= 1 \cdot \frac{1}{36} + 2 \cdot \frac{3}{36} + 3 \cdot \frac{5}{36} + 4 \cdot \frac{7}{36} + 5 \cdot \frac{9}{36} + 6 \cdot \frac{11}{36} = \frac{161}{36} = 4.472$$

9. Variance

La **variance** *Var(X)* d'une variable aléatoire *X* mesure **l'étendue ou la dispersion** d'une variable aléatoire autour de la "moyenne" *E(x).* Plus elle est élevée, plus les valeurs sont dispersées.

$$Var(X) = \left[x_1 - E(X)\right]^2 f(x_1) + \left[x_2 - E(X)\right]^2 f(x_2) + \ldots$$
$$= \sum_{i=1}^{n}\left[x_i - E(X)\right]^2 f(x_i) = \sum_{i=1}^{n}\left[x_i - \mu\right]^2 f(x_i) \qquad (9.1)$$

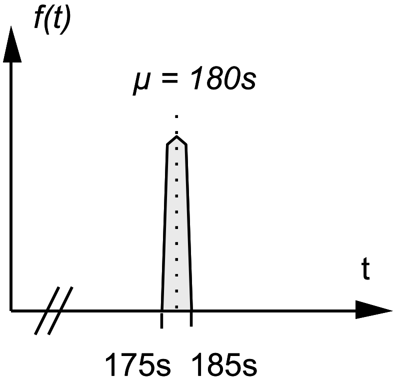
On montre facilement que la variance est aussi égale à

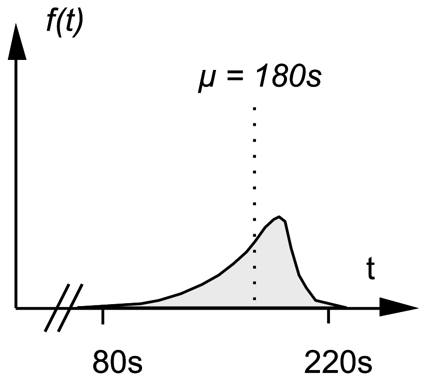
$$Var(X) = E\left(\left[X - E(X)\right]^2\right) = E(X^2) - E^2(X) \qquad (9.2)$$

Il est en général plus facile de calculer la variance à partir de l'expression (9.2) qu'à partir de la définition (9.1).

Illustration

On mesure la durée de transfert d'un grand nombre de transfert de fichiers multimédia dont la valeur moyenne est de 180s. Si les mesures sont entre 175s et 185s la variance est faible. Par contre, si les valeurs mesurées sont réparties régulièrement entre 80s et 220s, la variance est élevée.





La variance est **un nombre** obtenu pour une fonction dans un cas d'expérience donné, ce **n'est pas une fonction**.

Dans le cas d'une variable aléatoire discrète infinie, on peut directement généraliser

$$Var(X) = \sum_{i=1}^{\infty}\left[x_i - \mu\right]^2 f(x_i) \tag{9.3}$$

Dans le cas d'une variable aléatoire continue, il faut intégrer sur tout l'axe réel:

$$Var(X) = \int_{-\infty}^{\infty}(x - \mu)^2 f(x)\,dx \tag{9.4}$$

Exemple

On reprend l'exemple de la Section 8 avec les deux dés et le maximum des deux chiffres.

x_i =	*1*	*2*	*3*	*4*	*5*	*6*
$f(x_i)$=	*1/36*	*3/36*	*5/36*	*7/36*	*9/36*	*11/36*

On a déjà calculé que E(X) = μ = 4.472. On calcule alors la quantité

$$E(X^2) = \sum_{i=1}^{6} x_i^2 f(x_i)$$
$$= 1^2 \cdot \frac{1}{36} + 2^2 \cdot \frac{3}{36} + 3^2 \cdot \frac{5}{36} + 4^2 \cdot \frac{7}{36} + 5^2 \cdot \frac{9}{36} + 6^2 \cdot \frac{11}{36} = \frac{791}{36} = 21.97$$

d'où on tire la variance selon (9.2)

$$Var(X) = E(X^2) - \mu^2 = 21.972 - 20.001 = 1.971$$

10. Ecart-type

Étant donné une variable aléatoire *X* qui admet une variance *Var(X),* on appelle écart-type le nombre

$$\sigma(X) = \sqrt{Var(X)} \tag{10.1}$$

On voit toute suite que dans le cas d'une variable aléatoire constante *X* = Cte, *Var(X)* = 0 et *s(X)* = 0. On utilise souvent la notation $s^2(X) = Var(X)$

Exemple

Dans l'exemple de la Section 9, on obtient l'écart-type directement à partir de la variance

$$\sigma(X) = \sqrt{1.971} = 1.404$$

11. Références

Il existe de nombreux ouvrages dans le domaine des probabilités. La plupart des textes de télétrafic possèdent aussi un résumé sur les probabilités. Quelques titres sont donnés ci-dessous (en français, en allemand et en anglais):

Orientés pour les ingénieurs:

A. Ruegg, **Probabilités et statistique.** Presses polytechniques romandes, Lausanne, 4ème éd.,1994.

F. Barth et R. Hasler, **Stochastik**, Ehrenwirth, 1984.

G. Ecoffey, **Statistique,** EIA-FR, 1997.

M.A. Schnetzer, **Statistiques & Probabilités**, EIA-FR, 2006

S. Lipschutz, **Probabilités.** Schaum, McGraw-Hill, New York/Paris, 1973.

S. Ross, **A first course in probability.** MacMillan, New York, 1984.

Plus théoriques:

P. Louquet et A. Vogt, **Probabilités.** Armand Colin, Paris, 1971

Probabilités traitées dans le cadre du télétrafic:

L. Kleinrock, **Queuing systems**. 2 volumes. Wiley, New York, 1975.