Cours d'Éléments de Statistique

Jean Christophe Meunier

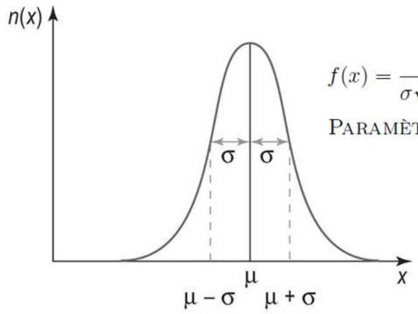
MODULE 6
Lois de distribution

1ère Bac, Commerce Extérieur
Année Académique 2015-2016

A. Introduction

II. Lois de distribution

- Donne la probabilité d'occurrence de toutes les valeurs d'une variable théorique
- Distribution caractérisée par :
  - Fonction : équation qui donne la forme
  - Paramètres : habituellement tendance centrale et dispersion
    - Ex. loi normale



$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

PARAMÈTRES : μ et σ

II. Lois de distribution

- Préalable à la compréhension des lois de distribution
  - Analyse combinatoire
  - Probabilité
- Notion essentielle : Variables aléatoires (v.a.)
  - Résultat d'une épreuve dans un phénomène aléatoire
  - Se traduit par une « grandeur » mathématique
  - Peut être discrète (nombre entier) ou continue (nombre réel)

B. Lois de probabilités de variables aléatoires (v.a.) discrètes

5

Introduction : v.a. discrètes

1. Définition d'une v.a. discrète
   - Soit Ω un espace fondamental et *P* une probabilité sur Ω. On appelle *variable aléatoire discrète* toute application X de Ω dans une partie de finie ou dénombrable de R

$$\begin{aligned} X \ : \ \Omega &\longrightarrow \mathbb{R} \qquad\qquad avec \qquad X(\Omega) \ \textit{fini ou dénombrable} \\ \omega &\longmapsto X(\omega) \end{aligned}$$

   - Etant donnée une variable aléatoire discrète X telle que $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$, on appelle *loi de probabilité* (ou distribution de probabilité) de X une expression des probabilités

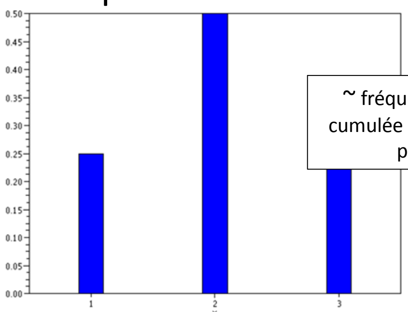
$$p_i = \mathbb{P}(X = x_i), \ \ i = 1 \ldots n$$

   - Les probabilités p_i trouvées vérifient alors $\sum_{i=1}^{n} p_i = 1$

6

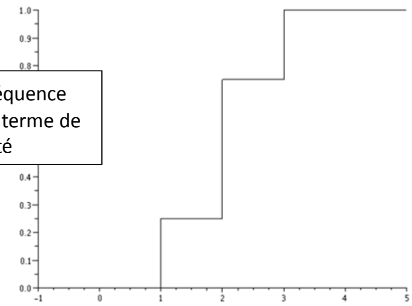
Introduction : v.a. discrètes

2. Illustration : lancé de deux pièces de monnaie
   - Ω comprend 4 e_i - notés PP, PF, FP, FF – chacun de probabilité 1/4.
   - v.a. discrète X = nombre de piles obtenus
     - P(X = 0) =1/4 ; P(X = 1) =1/2 ; P(X = 2) =1/4
3. Représentation d'une v.a. discrète



Densité de fréquences relatives

~ fréquence/fréquence cumulée mais en terme de probabilité



Fonction de répartition

Introduction : v.a. discrètes

4. Paramètres d'une v.a. : Espérance (~moyenne)
   - Lorsqu'on répète un grand nombre de fois une expérience, on peut estimer la moyenne des valeurs obtenues
   - Espérance mathématique : moyenne des valeurs prises par la v.a. X pondérées par la *P* d'obtenir cette valeur.
   - Soit X une v.a. discrète sur Ω telle que $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$ avec $x_1 < \ldots < x_n$. On appelle espérance mathématique de X, le réel

$$\mathbb{E}(X) = \sum_{i=1^n} x_i \mathbb{P}(X = x_i) = \sum_{i=1^n} x_i p_i$$

   - Ex. du lancer des deux pièces de monnaies

$$\mathbb{E}(X) = 0 \times \frac{1}{4} + 1 \times \frac{1}{2} + 2 \times \frac{1}{4} = 1$$

Introduction : v.a. discrètes

4. Paramètres d'une v.a. : Variance et écart type
   - dispersion de la distribution de probabilité
   - Soit X une v.a. discrète sur Ω telle que $X(\Omega) = \{x_1, x_2, \ldots, x_n\}$, avec $x_1 < \ldots < x_n$
     - Variance de X : Var(X) ou σ^2

$$Var(X) = \mathbb{E}\left[\left(X - \mathbb{E}(X)\right)^2\right] = \sum_{i=1}^{n} p_i \left(x_i - \mathbb{E}(X)\right)^2$$

     - Ecart-type de X : σ

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

9

I. Loi de Bernoulli

- Modèle :
  - On effectue <u>*un tirage*</u> d'une boule dans une urne contenant des boules blanches (x=1) et noires (x=0), les premières en proportion *p*, les secondes en proportion *q* = 1-*p*. Soit X le nombre de boules blanches tirées
- Notation :
  - $X \sim \mathcal{B}(1, p)$ avec paramètre *p*
- Valeurs possibles et probabilité de ces valeurs
  - Valeurs possibles : $X(\Omega) = \{0, 1\}$
  - Probabilités : $P(X = 0) = q = 1 - p$; $P(X = 1) = p$
- Espérance et variance
  - Espérance : $E(X) = p$
  - Variance : $Var(X) = pq$

10

II. Loi binomiale

- Modèle :
  - On effectue *n tirages* avec remise dans une urne contenant des boules blanches et noires, les premières en proportion *p*, les secondes en proportion *q* = 1-*p*. Soit X le nombre de boules blanches tirées
- Notation :
  - $X \sim \mathcal{B}(n,p)$ avec paramètres : p, n
- Valeurs possibles et probabilités
  - Valeurs possibles : X(Ω) = {0, 1, 2,…,n}
  - Probabilités : $\forall k \in (\Omega) \quad P(X=k) = C_n^k p^k q^{n-k}, \quad \text{où } C_n^k = \frac{n!}{k!\,(n-k)!}$
- Espérance et variance
  - Espérance : E(X) = *np*
  - Variance : Var(X) = *np(1-p) = npq*

Rappel :
formule combinaison SANS répétition

11

II. Loi binomiale

- Hypothèse de la loi binomiale
  - Type de réponse unique (succès) ou double mutuellement exclusive (boule blanche vs. boule noire)
  - Les résultats à chacun des essais de n sont indépendants l'un de l'autre
  - La P de 'succès' est constante pour chacun des essais

12

II. Loi binomiale

- Exemple
  - Une pièce de monnaie est truquée → à chaque lancer j'ai une chance sur trois de faire "pile".
    - Quelle est la probabilité en 4 lancers d'avoir une fois pile ?
  - Que connait-on ?
    - X = 'avoir pile'
    - *p* = probabilité d'avoir pile (sur un lancer) = 1/3 = .33 = 33%
    - *n* = nombre de lancers = 4
    - *k* = 'avoir 1 fois pile' (à l'issue des 4 lancers)
  - Application de formule

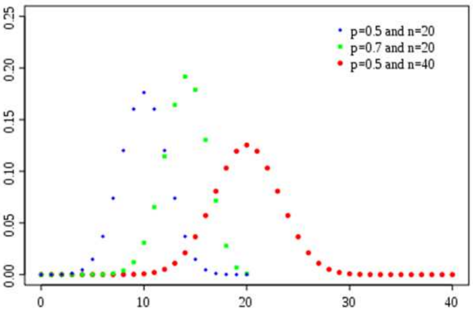
$$\forall k \in (\Omega) \quad P(X = k) = C_n^k p^k q^{n-k}, \quad \text{où } C_n^k = \frac{n!}{k!\,(n-k)!}$$

  - Calculs fastidieux → tableaux de probabilité (cf. IIIb)

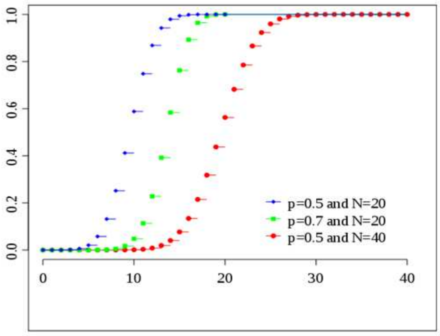
II. Loi binomiale

- Représentation

Densité de fréquences relatives



Fonction de répartition



II. Loi binomiale

- Propriétés
  - loi de Bernoulli = Loi binomiale pour n = 1
    - Binomiale : n tirages → Bernoulli : 1 tirage (n=1)

      $X \sim \mathcal{B}(n,p)$ $\quad\quad$ $X \sim \mathcal{B}(1,p)$
  - Si X_1 et X_2 sont des v.a. indépendantes telles que $X_1 \sim \mathcal{B}(n_1,p)$ et $X_2 \sim \mathcal{B}(n_2,p)$
    - alors, $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$
  - Si X_1, …, X_n sont des v.a. indépendantes suivant une loi de Bernoulli de paramètre *p*
    - alors, $X_1 + \cdots + X_n \sim \mathcal{B}(n,p)$

Exercice 1

Un examen consiste en un questionnaire à choix multiples de 12 questions, comportant chacune 5 réponses possibles. Pour être reçu, il faut répondre correctement à 8 questions au moins sur les 12. Quelle est la probabilité d'être reçu à l'examen :

a. Si on se présente sans avoir rien appris et si on répond au hasard ?

b. Si on a travaillé suffisamment pour pouvoir éliminer 3 réponses à chaque question, et si on choisit au hasard entre les 2 réponses restantes ?

c. Si on a travaillé suffisamment pour pouvoir répondre de façon *sûrement* exacte à deux questions. Pour les 10 questions restantes on choisit la réponse au hasard.

Formulation de chaque problème

X = 'réussite d'une question'

p = 'probabilité de réussite pour chaque question'

n = 'nombre de questions'

k = 'nombre de réussites' (sur n)

Exercice 1 : Solution

- A. n=12 ; k=8,9,10,11,12 ; p = 0.2
- B. n=12 ; k=8,9,10,11,12 ; p = 0.5
- C. n=10 ; k=6,7,8,9,10 ; p = 0.2

Exercice 1 : Solution

- Retour aux formules :

  Un examen consiste en un questionnaire à choix multiples de 12 questions, comportant chacune 5 réponses possibles. Pour être reçu, il faut répondre correctement à 8 questions au moins sur les 12. Quelle est la probabilité d'être reçu à l'examen :

  a. Si on se présente sans avoir rien appris et si on répond au hasard ?

  – Que connait-on ?
    - ***n***=12 ; ***k***=8,9,10,11,12 ; ***p*** = 0.2
    - Également *q*, sachant que ***q*** = 1-*p* = 0.8
    - Tout ce qu'il faut pour connaître probabilité, espérance et variance

Exercice 1 : Solution

- Retour aux formules :
  - *Probabilité*
    - ex. pour strictement 8 bonnes réponses sur 12

$$P(X = k) = C_n^k p^k q^{n-k} = \frac{n!}{k!\,(n-k)!}\; p^k \; q^{n-k}$$

Toutes les combinaisons possibles de 8 (k) bonnes réponses sur 12(n)
(Combinaison sans répétition car pas deux fois même question dans XM)

Probabilité de ***réussite*** (p=0,2) exposant nombre de ***réussites*** qui nous intéresse (k=8)

Probabilité d'***échec*** (p=0,8) exposant nombre d'***échecs*** qui nous intéresse (n-k=4)

- Sachant que *n*=12 ; *k*=8 ; *p* = 0.2 et *q* =0.8

$$P(X = 8) = C_{12}^8 0{,}2^8 0{,}8^{12-8} = \frac{12!}{8!\,(12-8)!}\; 0{,}2^8 \quad 0{,}8^{12-8}$$

Exercice 1 : Solution

- Retour aux formules :
    - Sachant que n=12 ; k=8 ; p = 0.2 et q =0.8
  - *Espérance* : E(X) = *np = 12 * 0,2 = 2,4*
    - En n'ayant rien étudié, sachant que l'on a 1/5 chance de réussir à chacune question :
      - **2,4/12** est le résultat moyen auquel on devrait s'attendre (~le plus probable)
  - *Variance* : σ^2 = Var(X) = *np*(1-*p*) = *npq = 12*0,2*0,8 = 1,92*
  - *Ecart-type* : σ = racine carré de 1,92 = 1,38
    - Ce résultat (2,4/12) serait susceptible de varier de 1,38 points
      - La plupart des étudiants ayant recourt à cette stratégie auraient des points allant de **1,02/12** (2,4-1,38) à **3,78/12** (2,4+1,38)

Exercice 2

On extrait au hasard un échantillon de 5 électeurs d'une vaste population où la proportion des républicains est de 60%.

Le nombre de républicains dans l'échantillon peut varier entre 0 et 5.

Donner la distribution de probabilité de cette variable aléatoire

Calculer la moyenne et l'écart type

- Que connait-on ?
  - X = 'obtenir un républicain'
  - *n* = 'nombre de sujets tirés'
  - *p* = 'probabilité d'obtenir un républicain' sur un tirage
- Rem : distribution de probabilité ?
  - Probabilité pour chacune des valeurs possible de k (c'est à dire 0,1,2,3,4,5)

Exercice 2 : Solution

(a) Distribution		(b) Moyenne	Variance		
s	$p(s)$	$sp(s)$	$s-\mu$	$(s-\mu)^2$	$(s-\mu)^2\,p(s)$
0	.010	0	−3	9	.090
1	.077	.077	−2	4	.308
2	.230	.460	−1	1	.230
3	.346	1.038	0	0	0
4	.259	1.036	1	1	.259
5	.078	.390	2	4	.312
	1.000	E(x) = 3.00			$\sigma^2 = 1.20$
					$\sigma = 1.10$

$\boxed{\sigma^2 = npq = 5 * 0,6 * 0.4 = 1,2}$

$\boxed{E(X) = np = 5 * 0,6 = 3}$

table IIIb

III. Loi de Poisson

- Modèle :
  - Soit X, le nombre d'apparitions d'un événement rare
    - Rare = sur un intervalle 'limité' (laps de temps ou région de l'espace)
- Notation
  - $X \sim \mathcal{P}(\lambda)$ avec paramètre λ
    - λ est la moyenne du nombre d'occurrence dans l'intervalle considéré
    - Ex : évènement se produit en moyenne 4 x/min, pour étudier le nombre d'évènements se produisant dans un laps de temps de 10 minutes → loi de Poisson de paramètre λ = 10 × 4 = 40

23

III. Loi de Poisson

- Valeurs possibles et probabilités
  - Valeurs possibles : X(Ω) = N
  - Probabilités :

$$\forall\, k \in X\,(\Omega) \qquad P(X = k) = e^{-\lambda} . \frac{\lambda^k}{k!}$$

Voir Tableau X

e = nombre népérien = 2,71828

- Espérance et variance
  - Espérance : E(X) = λ
  - Variance : Var(X) = λ

24

III. Loi de Poisson

- Exemples de v.a. de Poisson
  – Nombre de voitures arrivant dans une station sur un intervalle d'une heure
  – Nombre de défauts de paiement (mensualités) pour abonnement annuel à un service
  – Nombre d'appareils défectueux dans une chaine de production de 1.000 unités/jour
- Hypothèse de base de loi de Poisson
  1. Nombre d'occurrences intervalle x indépendant du nombre d'occurrences dans intervalle y
  2. *P* d'occurrences identiques pour tout intervalle de même taille
  3. *P* d'occurrences proportionnelles à la taille de l'intervalle

25

III. Loi de Poisson

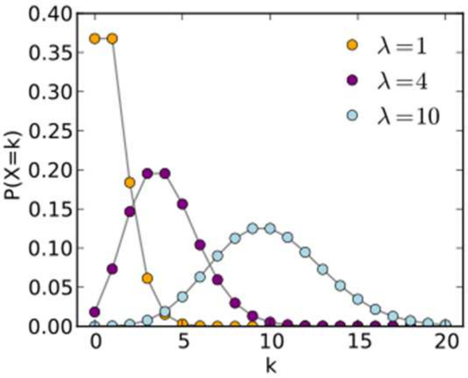
- Propriété
  – Si X_1 et X_2 sont des v.a. indépendantes telles que $X_1 \sim \mathcal{P}(\lambda_1)$ et $X_2 \sim \mathcal{P}(\lambda_2)$, alors $X_1 + X_2 \sim \mathcal{P}(\lambda_1 + \lambda_2)$
- Binomiale → Poisson
  – Lorsque *n* tend vers l'infini et que *p* tend vers 0, la loi binomiale converge vers loi de Poisson avec paramètre $\lambda = np$
  – *En pratique, Binomiale → Poisson*
    - Dès que $n > 30$ et $np < 5$
    - ou, dès que $n > 50$ et $p < 0{,}1$

26

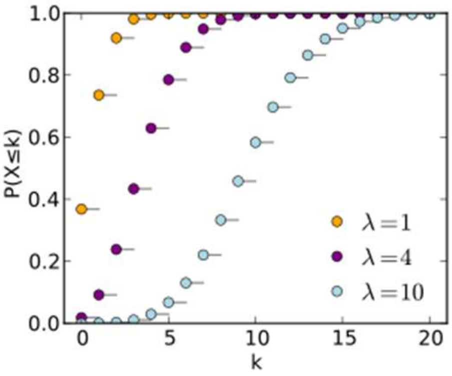
III. Loi de Poisson

- Représentation

Densité de fréquences relatives



Fonction de répartition



27

Exercice

Le nombre d'ordinateurs vendus chaque jour dans un magasin spécialisé suit une loi de Poisson de paramètre 4.

Calculer la probabilité dans une journée :

1. De ne vendre aucun ordinateur
2. De vendre 4 ordinateurs
3. De vendre au moins un ordinateur
4. Le nombre d'ordinateurs vendus est compris entre 2 et 6

28

Exercice : Solution

1. aucun ordinateur : *p*(X=0)

   0,0183 = 1,83 %

2. quatre ordinateurs : *p*(X=4)

   0,1954 = 19,54%

3. au moins 1 ordinateur : *p*(X≥1)

   Chercher l'événement complémentaire : P (Ā) = aucun ordinateur = 0,0183 = 1,83 %

   P (A) = P(au moins 1) = 1-*P*(Ā=zéro ordinateur) = 1 – 0,0183 = 0,9817 = 98,17%

4. entre 2 et 6 ordinateurs : *p*(2≤X≤6)

   p(X=2) + *p*(X=3) + *p*(X=4) + *p*(X=5) + *p*(X=6) = 0,7978 = 79,78%

29

Exercice : Solution

Retour aux formules

- Ce qu'on connait : λ
  - Seul paramètre d'un distribution de Poisson
  - Moyenne du nombre d'occurrence dans l'intervalle considéré
- *Probabilité* de k=3 sachant λ=4
  – Probabilité de vendre 3 ordinateurs sur une journée sachant qu'en moyenne on en vend 4 par jour

$$P(X = k) = e^{-\lambda}.\frac{\lambda^k}{k!} = 2{,}718^{-4}.\frac{4^3}{3!}$$

- *Espérance* : E(X) = λ = 4
- *Variance* : σ² = Var(X) = λ = 4
- *Ecart-type* : σ = racine carrée de 4 = 2

30

c. Lois de probabilités de variables aléatoires (v.a.) continues

Introduction : v.a. continues

- Définition.
  - Soit Ω un espace fondamental et *P* une probabilité sur Ω. On note F la fonction de répartition de X définie par :

$$F \ : \ \mathbb{R} \longrightarrow [0,1]$$
$$x \longmapsto \mathbb{P}(X \leq x)$$

  - On dit que X est une variable aléatoire continue s'il existe une fonction f définie sur R telle que

(i) pour tout $x \in \mathbb{R}$, $\quad f(x) \geq 0$,

(ii) f est continue sur $\mathbb{R}$ sauf peut être en un nombre fini de points où elle admet une limite à droite et une limite à gauche finies,

(iii) $\displaystyle\int_{-\infty}^{+\infty} f(t)dt = 1$,

(iv) pour tout $x \in \mathbb{R}$, $\quad F(x) = \displaystyle\int_{-\infty}^{t} f(t)dt.$

Caractéristiques d’une v.a. continue

1. Espérance (~moyenne)
   – $$\mathbb{E}(X) = \int_{-\infty}^{+\infty} t f(t) dt$$

2. Variance (Var(x) ou σ^2) et écart-type (σ)

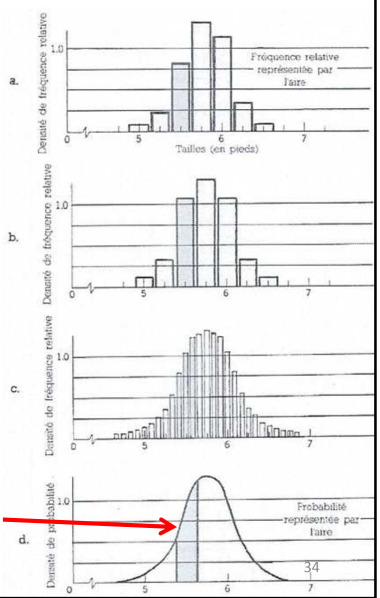
$$Var(X) = \mathbb{E}\left[(X - \mathbb{E}(X))^2\right] = \int_{-\infty}^{+\infty} (t - \mathbb{E}(X))^2 f(t) dt,$$
$$= \mathbb{E}(X^2) - (\mathbb{E}(X))^2 = \left(\int_{-\infty}^{+\infty} t^2 f(t) dt\right) - (\mathbb{E}(X))^2$$

$$\sigma(X) = \sqrt{Var(X)}.$$

33

Représentation d’une v.a. continue

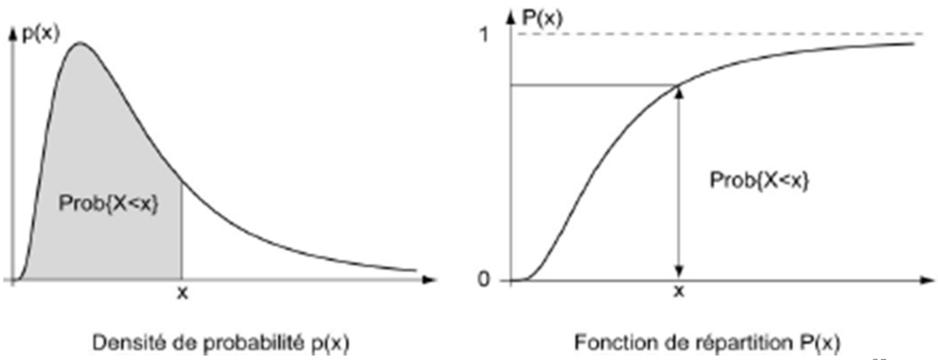
- De v.a. discrète à continue
  - Si n augmente et si amplitude des classes diminue
    - La fréquence tend vers la probabilité
    - La courbe ‘densité de fréquences relatives’ tend vers la ‘densité de probabilité’
  - En découle,
    - La probabilité d’un point est nulle car l’aire correspondante est un rectangle de largeur nulle
    - La probabilité qu’un événement X situé entre 2 valeurs est non nulle
      - aire sous la courbe entre ces deux valeurs
    - Aire totale sous la courbe = 1



34

Représentation d'une v.a. continue

Densité de probabilité *Fonction de répartition*



35

I. Loi uniforme

- Modèle (densité de probabilité)
  – une variable aléatoire continue X suit la loi uniforme sur l'intervalle [a, b] si elle admet pour densité la fonction f définie sur R par

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{si } x \in [a,b], \\ 0 & \text{sinon.} \end{cases}$$

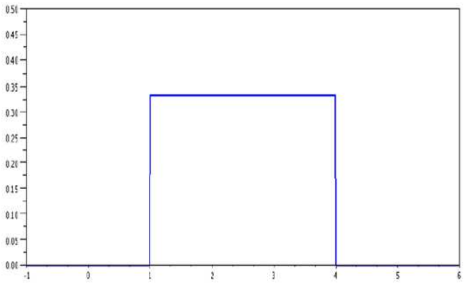
- Notation
  – X ~ U([a,b]) avec pour paramètres, les bornes a et b de l'intervalle
- Espérance et variance
  – Espérance : $\mathbb{E}(X) = \dfrac{a+b}{2}$
  – Variance : $Var(X) = \dfrac{(b-a)^2}{12}$

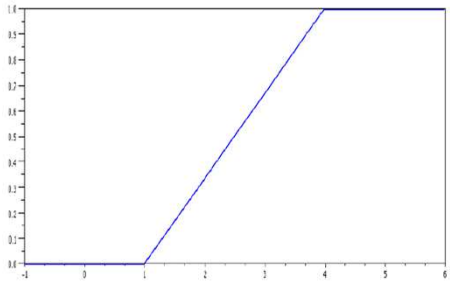
36

I. Loi uniforme

- Représentation : U([1,4])

Densité de probabilité *Fonction de répartition*





37

II. Loi exponentielle

- Modèle (densité de probabilité)
  - on se place dans un phénomène d'attente et on s'intéresse à la variable aléatoire qui représente le temps d'attente entre deux événements successifs ou encore une durée de vie
  - une variable aléatoire continue X suit la loi exponentielle de paramètre λ ($\lambda > 0$) si elle admet pour densité la fonction f définie sur R par

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

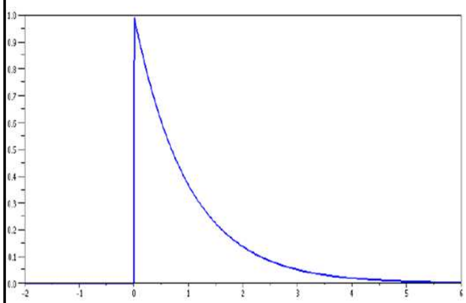
- Notation
  - $X \sim \mathcal{E}(\lambda)$ avec λ comme paramètre
- Espérance et variance
  - Espérance : $\mathbb{E}(X) = \frac{1}{\lambda}$
  - Variance $Var(X) = \frac{1}{\lambda^2}$

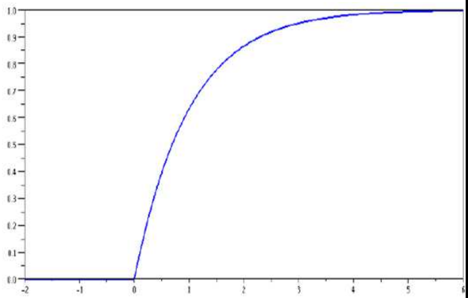
38

II. Loi exponentielle

- Représentation : $\varepsilon(\lambda)$

Densité de probabilité | *Fonction de répartition*





39

III. Loi normale

- Modèle (densité de probabilité)
  - une variable aléatoire continue X suit la loi normale de paramètres et (> 0) si elle admet pour densité la fonction f définie sur R par

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

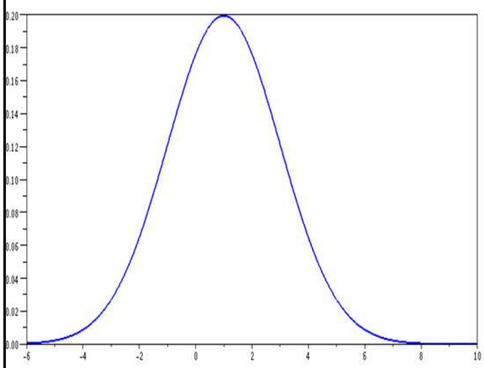
- Notation :
  - X ~ $\mathcal{N}(\mu,\sigma)$, avec μ et σ comme paramètre
- Espérance et variance :
  - Espérance : E(X) = μ
  - Variance : Var(X) = σ

40

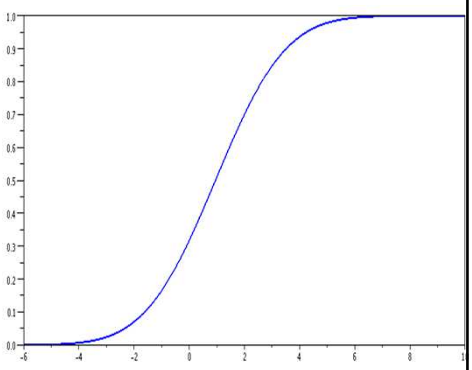
III. Loi normale

- Représentation : $\mathcal{N}(\mu,\sigma)$, avec $\mu = 1$, $\sigma = 2$

Densité de probabilité



Fonction de répartition



41

III. Loi normale

- Retour aux formules :
  - Sachant $\mathcal{N}(\mu,\sigma)$, avec $\mu = 1$, $\sigma = 2$
- Probabilité d'avoir X = 1,5

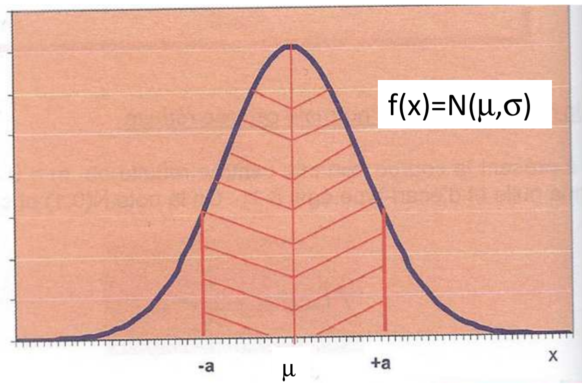
$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \longrightarrow P(X = 1{,}5) = f(1{,}5) = \frac{1}{2\sqrt{2\pi}}.e^{-\frac{1}{2}\left(\frac{1{,}5-1}{2}\right)^2}$$

- Espérance → Moyenne : $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$
- Variance : $\sigma^2 = \frac{\sum\left(X-\overline{X}\right)^2}{n}$
- Ecart-type : $\sigma = \sqrt{\frac{\sum\left(X-\overline{X}\right)^2}{n-1}}$

42

III. Loi normale : propriétés

- la distribution est symétrique par rapport à la droite $x = \mu$
  - Dès lors, pour tout a, $P(\mu-a<x<\mu) = P(\mu<x<\mu+a)$



43

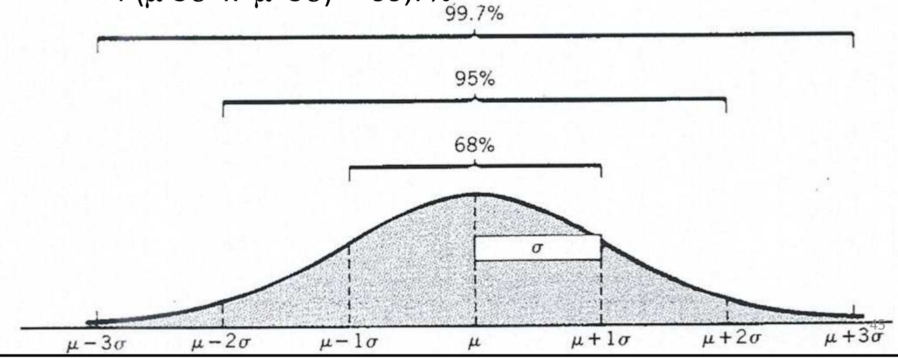
III. Loi normale : propriétés

- Additivité
  - Si $X_1 \sim N(\mu_1, \sigma_1)$, $X_2 \sim N(\mu_2; \sigma_2)$, et si X_1 et X_2 sont indépendantes, alors $X_1 + X_2 \sim N(\mu_1 + \mu_2, \sigma_1 + \sigma_2)$

44

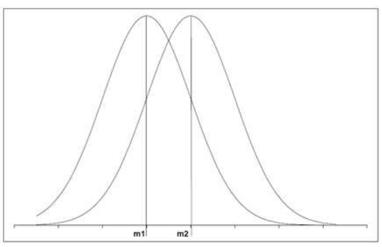
III. Loi normale : propriétés

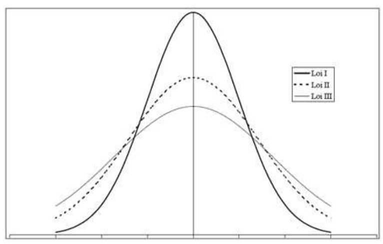
- Quelque soit la courbe normale de paramètre (μ,σ), l'espace de probabilité défini à partir des points μ et/ou σ reste le même, tel que :
  - $P(\mu-\sigma<x<\mu+\sigma) = 68\ \%$
  - $P(\mu-2\sigma<x<\mu+2\sigma) = 95\ \%$
  - $P(\mu-3\sigma<x<\mu+3\sigma) = 99{,}7\%$



III. Loi normale : propriétés

- Autant de courbes normales que de couples de paramètres (μ, σ)
  - μ définit la position
  - σ définit la dispersion (d'autant plus étalée que σ est grand)





- Dès lors, impossible d'avoir des tables de probabilité pour toutes les courbes normales possibles
  - Solution : ***<u>loi normale centrée réduite</u>***

IV. Loi normale centrée réduite: Loi de z

- Modèle (densité de probabilité)
  - Il s'agit de la loi normale obtenue pour μ= 0 et σ= 1. Elle admet pour densité la fonction f définie sur R par

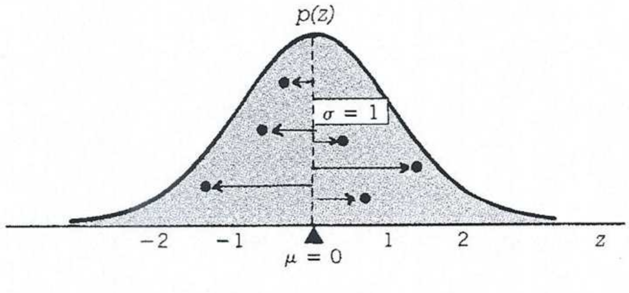
$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

- Notation :
  - Z ~ $N(0,1)$, avec $\mu=0$ et $\sigma=1$ comme paramètre
- Espérance et variance :
  - Espérance : E(X) = $\mu=0$
  - Variance : Var(X) = $\sigma=1$

47

IV. Loi normale centrée réduite : Score-z

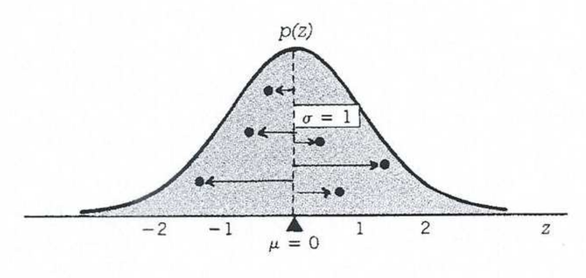
- Transformation des score brute (x) en score centré réduit (score-z)
  - Centrage: $x_i - \mu$
  - Réduction: diviser par σ
  - Données centrées réduites: $z_i = (x_i - \mu)/\sigma$



48

IV. Loi normale centrée réduite : Score-z

- Quelque soit la loi normale d'origine (de paramètre μ et σ)
  - la loi centrée réduite correspondante est toujours $N(0,1)$
  - Chaque réalisation de Z (score-z) est la distance à la moyenne en 'unité' d'écart-type
    - Ex : $z = 0 = \mu$; $z = 1 = \mu+\sigma$; $z=1,5 = \mu + 1,5\sigma$



49

IV. Loi normale centrée réduite : Score-z

- Le centrage et la réduction = transformation linéaire
  - permet d'exprimer des variables différentes sur une échelle commune, en les débarrassant de leurs unités physiques

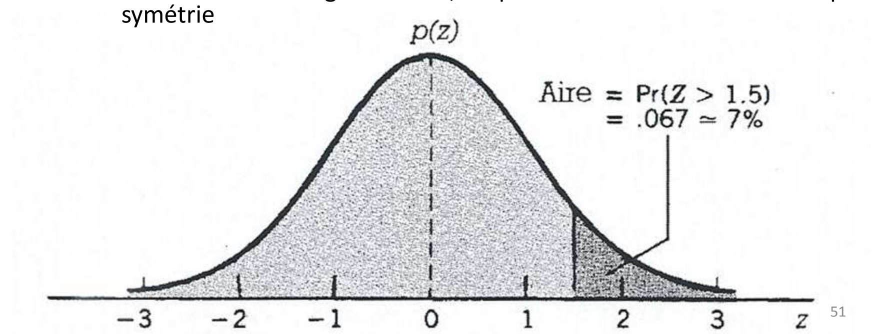
50

IV. Loi normale centrée réduite : Score-z

- L'échelle unique (en score-z) permet de recourir à une seule et même table de probabilités pour toutes courbes $N(\mu,\sigma)$

Table IV

- permet de trouver une valeur supérieure à la valeur centrée réduite
- Pour les valeurs négatives de z, les probabilités sont déterminées par symétrie

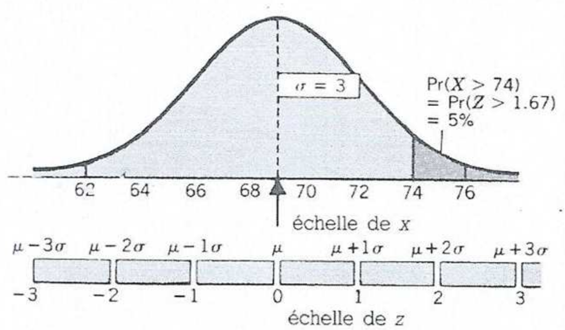


51

IV. Loi normale centrée réduite : Score-z

- Illustration :

Si on étudie la taille (exprimée en pouces, pour vous en donner un petit coup !) de la totalité des hommes américains et que l'on représente le caractère par une distribution des fréquences , on obtient la distribution normale suivante, de moyenne μ = 69 pouces et d'écart type σ = 3 pouces.



52

Comment calculer la proportion d'hommes dont la taille dépasse 74 pouces ?

IV. Loi normale centrée réduite : Score-z

- Illustration (suite) :

La figure montre qu'une taille de 74 pouces dépasse la moyenne de 2 écart types environ.

Pour calculer exactement l'écart à la moyenne, c'est-à-dire Z, on procède en 3 étapes:

1. Première étape: calculer l'écart à la moyenne (X – μ)
   On calcule l'écart de la valeur particulière (74 pouces) à la moyenne μ:
   Ecart à la moyenne = 74 – 69 = 5 pouces
2. Deuxième étape: comparer cet écart (X – μ) avec σ, c'est-à-dire calculer Z
   Combien cet écart représente-t-il d'écart types ?
   σ = 3 pouces et donc Z = 5/3 = 1,67 écart type

   $$\boxed{Z = \frac{x-\mu}{\sigma}}$$
3. Utiliser la table IV fournie en annexe pour obtenir la probabilité cherchée
   Pr(Z>1,67) = 0,047 soit 5% arrondi

53

Exercice 1

Soit Z la variable normale centrée réduite. Calculer:

a) Pr (Z>1,64)
b) Pr (Z<-1,64)
c) Pr (1,0 < Z < 1,5)
d) Pr (-1 < Z < 2)
e) Pr (-2 < Z < 2)

54

Exercice 1 : Solution

Solution

En utilisant la table IV de l'annexe, on peut calculer chaque probabilité.

a) Pr (Z>1,64) = 0,051 soit 5% arrondi

b) Par symétrie, Pr (Z<-1,64) = Pr (Z>1,64) = 0,051

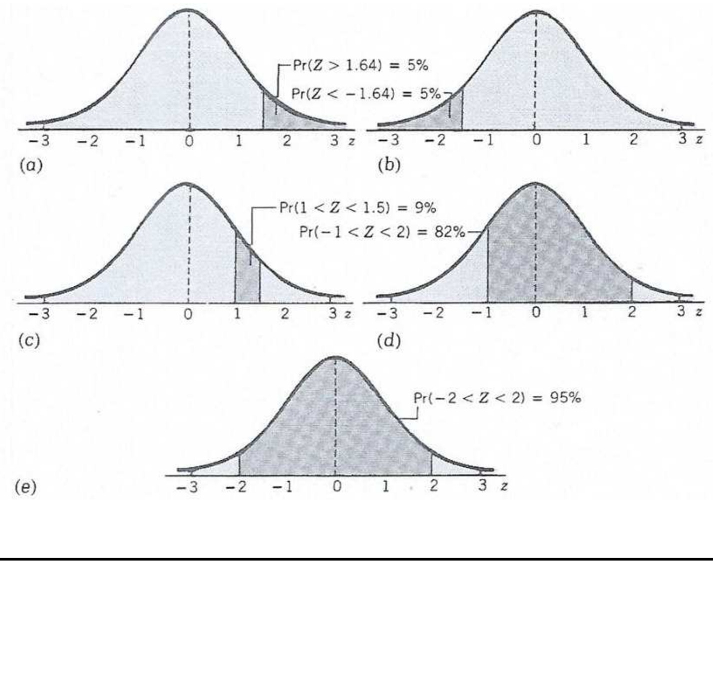
c) Pr (1,0 < Z < 1,5) = Pr (Z>1,0) – Pr (Z>1,5) = 0,159 – 0,067 = 0,092 soit 9% arrondi

d) Retrancher la surface des 2 queues de distribution de l'aire unitaire totale

Pr (-1 < Z < 2) = 1 – Pr(Z<-1) – Pr (Z>2) = 1 – 0,159 – 0,023 = 0,818

e) Pr (-2 < Z < 2) = 1 – Pr(Z<-2) – Pr(Z>2) = 0,95

Exercice 1 : Solution



Exercice 2

Soit X une variable aléatoire normale de moyenne $\mu = 16$ et d'écart type $\sigma = 5$. Calculer

a. $\Pr(X > 20)$

b. $\Pr(20 < X < 25)$

c. $\Pr(X < 10)$

d. $\Pr(12 < X < 24)$

Rem : Opération deux étapes

1. On commence par calculer z $Z = \dfrac{x-\mu}{\sigma}$
2. Une fois z connu, consulter table IV pour obtenir la probabilité correspondante

Exercice 2 : Solution

Solution

a) On commence par calculer Z: $Z = \dfrac{x-\mu}{\sigma} = \dfrac{20-16}{5} = 0,8$

Pr(X>20) = Pr(Z>0,8). En consultant les tables IV, on détermine cette probabilité.

b) On calcule Z pour X = 20 et pour X = 25

Ensuite, on détermine les valeurs de la probabilité correspondante dans les tables IV.

c) Idem

d) Idem

Exercice 3

Phil et Kim Bell ne savent pas s'ils doivent acheter une maison maintenant ou bien attrendre encore un an, auquel cas l'augmentation des prix risque de mettre cet achat hors de leur portée. Ils estiment que, s'ils attendent encore un an, l'augmentation des prix pourrait être normale avec une moyenne de 8 % et, compte tenu de l'incertitude du marché, un écart type de 10 %.

a. Si l'augmentation des prix excède 25%, ils savent qu'ils ne seront plus en mesure d'acheter. Quelle est la probabilité de cet événement ?

b. D'un autre côté, une baisse des prix rendrait leur attente très profitable. Quelle est la probabilité de cet événement ?

Exercice 3 : Solution

Solution

a) On doit calculer Pr(X>25). La distribution étant normale, on calcule Z

$Z = \frac{x-\mu}{\sigma} = \frac{25-8}{10} = 0,13$ et grâce aux tables IV, on détermine Pr(Z>0,13) = 0,448

b) Une baisse de prix équivaut à un pourcentage négatif en terme d'évolution, à savoir < 0%. On doit donc calculer Pr(X<0). La distribution étant normale on calcule Z

Z = (x-μ)/σ = (0 – 8)/10 = -0,8

Parce que la loi normale est symétrique, on sait que Pr(Z< -0,8) = Pr(Z> 0,8)

Grâce aux tables IV, on détermine que Pr(Z < - 0,8) = 0212 = 21,2 %