

SCI03 - - Analyse de données expérimentales

Rappels de probabilités

Thierry Denœux¹

¹Université de Technologie de Compiègne
tél : 44 96
tdenoeux@hds.utc.fr

Automne 2014

Expérience aléatoire

- On appelle **expérience aléatoire** une expérience qui, répétée plusieurs fois dans des conditions opératoires identiques, produit des résultats qui peuvent être différents.
- Les exemples classiques sont issus des jeux de hasard : lancer d'un ou plusieurs dés, tirage dans une urne, etc.

Modélisation

- Mathématiquement, la notion d'expérience aléatoire \mathcal{E} se formalise en définissant :
 - 1 un **ensemble fondamental** Ω définissant l'ensemble des résultats possibles de \mathcal{E} , appelés **événements élémentaires** ;
 - 2 un ensemble \mathcal{A} de parties de Ω (**événements**), clos par complémentation et union dénombrable (**tribu** ou **σ -algèbre**) ;
 - 3 une fonction $\mathbb{P} : \mathcal{A} \rightarrow [0, 1]$, appelée **mesure** ou **distribution** de probabilité, qui à tout événement A associe un nombre $\mathbb{P}(A)$ appelé probabilité de cet événement.

Probabilité

Définition

- $\mathbb{P}(\Omega) = 1$;
- Pour toute famille $(A_i)_{i \in I}$ finie ou dénombrable d'événements deux à deux disjoints $(A_i \cap A_j = \emptyset, \forall i, j \in I)$,

$$\mathbb{P} \left(\bigcup_{i \in I} A_i \right) = \sum_{i \in I} \mathbb{P}(A_i)$$

(additivité).

- La structure (Ω, \mathcal{A}) est appelée **espace probabilisable**, et $(\Omega, \mathcal{A}, \mathbb{P})$ **espace probabilisé**.

Probabilité

Interprétation

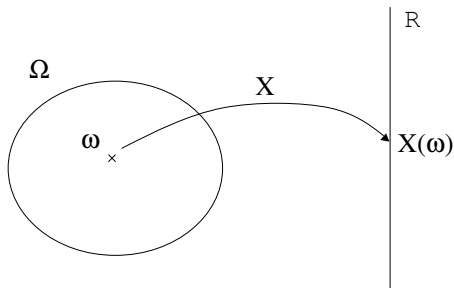
- $\mathbb{P}(A)$ s'interprète comme la **limite de la fréquence** de réalisation de l'événement A , lorsque le nombre n de répétitions de l'expérience aléatoire tend vers l'infini :

$$\mathbb{P}(A) = \lim_{n \rightarrow \infty} \frac{n_A}{n},$$

n_A étant le nombre de réalisations de l'événement A obtenu au cours de n répétitions de \mathcal{E} .

Définition

- On appelle **variable aléatoire (réelle)** une grandeur numérique dont la valeur est fonction du résultat d'une expérience aléatoire.

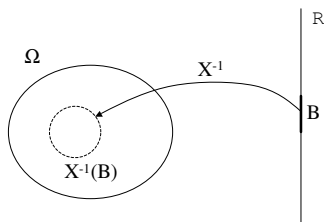


Loi de probabilité

- Soit B un intervalle de \mathbb{R} . On peut définir la probabilité que la v.a. X prenne sa valeur dans B comme

$$\mathbb{P}_X(B) = \mathbb{P}(\{\omega \in \Omega \mid X(\omega) \in B\}) = \mathbb{P}(X^{-1}(B)),$$

(On suppose que $X^{-1}(B) \in \mathcal{A}$ pour tout B).



- La donnée de $\mathbb{P}_X(B)$ pour tout intervalle B définit la **loi (ou distribution) de probabilité** de X .

Fonction de répartition

- Pour décrire complètement \mathbb{P}_X , il suffit de donner la **fonction de répartition** de X

$$F_X : \mathbb{R} \rightarrow [0, 1] \\ x \rightarrow \mathbb{P}_X(] - \infty, x]),$$

ce que l'on note $F_X(x) = \mathbb{P}(X \leq x)$.

- Propriétés suivantes :

- 1 croissance ;
- 2 continuité à droite : $\lim_{\epsilon \rightarrow 0, \epsilon > 0} F_X(x + \epsilon) = F_X(x)$;
- 3 conditions aux limites :
 $\lim_{x \rightarrow -\infty} F_X(x) = 0, \quad \lim_{x \rightarrow \infty} F_X(x) = 1.$

Fonction de probabilité (v.a. discrète)

- Une loi de probabilité peut également être définie par la fonction qui à chaque élément de V_X associe sa probabilité (loi discrète), ou par un équivalent dans le cas continu.
- Soit X une v.a. discrète. On appelle **fonction de probabilité** de X la fonction :

$$\begin{aligned} p_X : \mathbb{R} &\rightarrow [0, 1] \\ x &\rightarrow \mathbb{P}_X(\{x\}). \end{aligned}$$

- Propriétés :

$$\forall B \in \mathcal{B}(\mathbb{R}), \quad \mathbb{P}_X(B) = \sum_{x \in B} p_X(x)$$

$$\sum_{x \in V_X} p_X(x) = \mathbb{P}(V_X) = 1.$$

Fonction de densité de probabilité (v.a. continue)

- Dans le cas continu, la notion de **fonction de densité de probabilité** remplace celle de fonction de probabilité.
- Une v.a. continue X est dite **absolument continue** s'il existe une fonction $f_X : \mathbb{R} \rightarrow \mathbb{R}_+$, (**densité**) t.q.

$$\forall B \in \mathcal{B}(\mathbb{R}), \quad \mathbb{P}_X(B) = \int_B f_X(t) dt.$$

- Propriétés :

$$\int_{-\infty}^{+\infty} f_X(t) dt = \mathbb{P}_X(\mathbb{R}) = 1 \quad \text{et} \quad F_X(x) = \int_{-\infty}^x f_X(t) dt.$$

- La quantité $f_X(x)$ n'est pas une probabilité. C'est l'intégrale de f_X sur un intervalle qui est une probabilité.

Loi de Bernoulli

- Soit une expérience aléatoire \mathcal{E} , et A un événement associé à \mathcal{E} , de probabilité p .
- Soit X la fonction indicatrice de A , définie pour tout $\omega \in \Omega$ par :

$$X(\omega) = \begin{cases} 1 & \omega \in A \\ 0 & \text{sinon.} \end{cases}$$

- Fonction de probabilité :

$$p_X(1) = \mathbb{P}(A) = p, \quad p_X(0) = \mathbb{P}(\bar{A}) = 1 - p$$

- Par définition, on dit que X suit une **loi de Bernoulli de paramètre p** , ce que l'on note $X \sim \mathcal{B}(p)$.

Loi binomiale

- Supposons que l'on répète n fois l'expérience précédente, de manière indépendante. On définit n v.a. de Bernoulli $X_i \sim \mathcal{B}(p)$, $i = 1, \dots, n$.
- Le nombre de réalisations de l'événement A est

$$Y = \sum_{i=1}^n X_i,$$

- Par définition, $Y \sim \mathcal{B}(n, p)$. Le domaine de définition est $V_X = \{0, \dots, n\}$.
- Fonction de probabilité de Y :

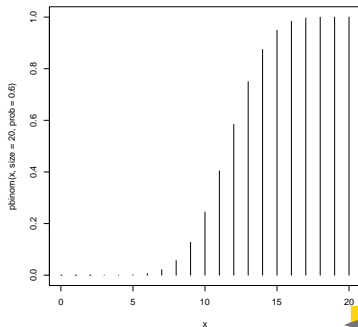
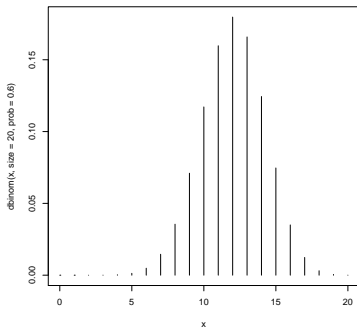
$$p_Y(y) = C_n^y p^y (1-p)^{n-y}, \quad \forall y \in \{0, \dots, n\}.$$

Loi binomiale en R

- Fonction de probabilité : `dbinom(x, size, prob)`.
- Fonction de répartition : `pbinom(q, size, prob)`.

```
x<-0:20
```

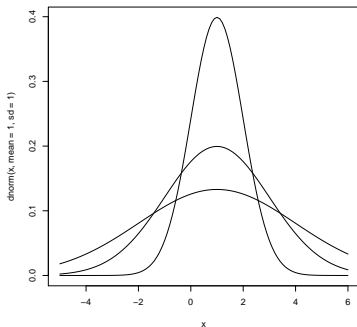
```
plot(x,dbinom(x,size=20,prob=0.6),type='h')
```



Loi normale

- La loi normale $\mathcal{N}(\mu, \sigma^2)$ est définie par la fonction de densité de probabilité :

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left[-\frac{1}{2} \left(\frac{x - \mu}{\sigma} \right)^2 \right],$$



Loi normale

Fonction de répartition

- La fonction de répartition correspondante n'a pas d'expression analytique.
- On l'exprime communément à l'aide de la fonction ϕ (fonction de répartition de la loi $\mathcal{N}(0, 1)$, appelée loi normale centrée-réduite) :

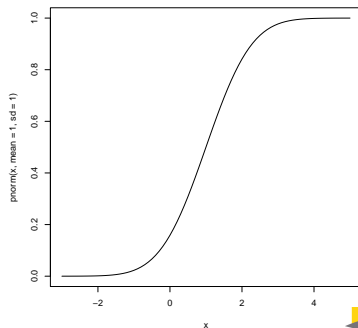
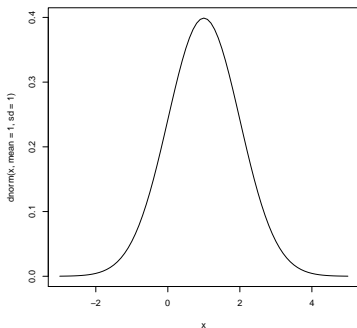
$$X \sim \mathcal{N}(\mu, \sigma^2) \Leftrightarrow F_X(x) = \phi\left(\frac{x - \mu}{\sigma}\right), \forall x \in \mathbb{R}.$$

- Propriété : $\phi(-x) = 1 - \phi(x)$ pour tout $x \in \mathbb{R}$.

Loi normale en R

- Fonction de probabilité : `dnorm(x, mean, sd)`.
- Fonction de répartition : `pnorm(q, mean, sd)`.

```
x<-seq(-3,5,0.1)  
plot(x,dbinom(x,mean=1,sd=1),type='l')
```



Espérance mathématique

Définition

- L'**espérance (mathématique)** d'une variable aléatoire réelle représente la « valeur moyenne » prise par cette variable aléatoire.
- Définition :

$$\mathbb{E}(X) = \begin{cases} \sum_{x \in V_X} xp_X(x) & \text{si } X \text{ est une v.a. discrète,} \\ \int_{\mathbb{R}} xf_X(x)dx & \text{si } X \text{ est une v.a. continue} \end{cases}$$

si ces quantités existent.

Espérance mathématique

Propriétés

- Si le graphe de la fonction de densité est symétrique par rapport à une valeur a (c'est-à-dire si $f_X(a-x) = f_X(a+x)$ pour tout x), alors $\mathbb{E}(X) = a$.
- Propriétés :
 - $\forall \alpha, \beta \in \mathbb{R}, \mathbb{E}(\alpha X + \beta) = \alpha \mathbb{E}(X) + \beta$.
 - Soit X une v.a. et φ une fonction $\mathbb{R} \rightarrow \mathbb{R}$. On a :

$$\mathbb{E}(\varphi(X)) = \begin{cases} \int_{\mathbb{R}} \varphi(x) f_X(x) dx & \text{si } X \text{ est une v.a. continue,} \\ \sum_{x \in V_X} \varphi(x) p_X(x) & \text{si } X \text{ est une v.a. discrète.} \end{cases}$$

Variance

Définition

- C'est une mesure de **dispersion** de la v.a. autour de son espérance :

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

- La racine carrée de la variance est appelée **écart-type** de la v.a. X et notée σ .

Variance

Propriétés

- $\text{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$.
- $\forall \alpha, \beta \in \mathbb{R}, \text{Var}(\alpha X + \beta) = \alpha^2 \text{Var}(X)$.
- Inégalité de Bienaymé-Tchebycheff :

$$\mathbb{P}(|X - \mathbb{E}(X)| \geq \varepsilon) \leq \frac{1}{\varepsilon^2} \text{Var}(X).$$

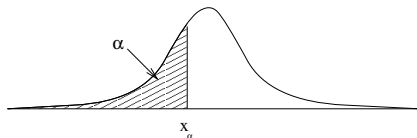
(Plus la variance est petite, plus faible est la probabilité que X s'écarte de son espérance d'une valeur donnée.)

Fractile

- Soit X une v.a. continue de fonction de répartition F_X .
- Pour tout $\alpha \in]0, 1[$, on appelle **fractile (quantile) d'ordre α** de X la quantité

$$x_\alpha = F_X^{-1}(\alpha).$$

- On a donc, par définition, $F_X(x_\alpha) = \alpha$, ou encore $\int_{-\infty}^{x_\alpha} f_X(x) dx = \alpha$.



- Fractiles de la loi normale en \mathbb{R} :
`qnorm(p, mean = 0, sd = 1)`.