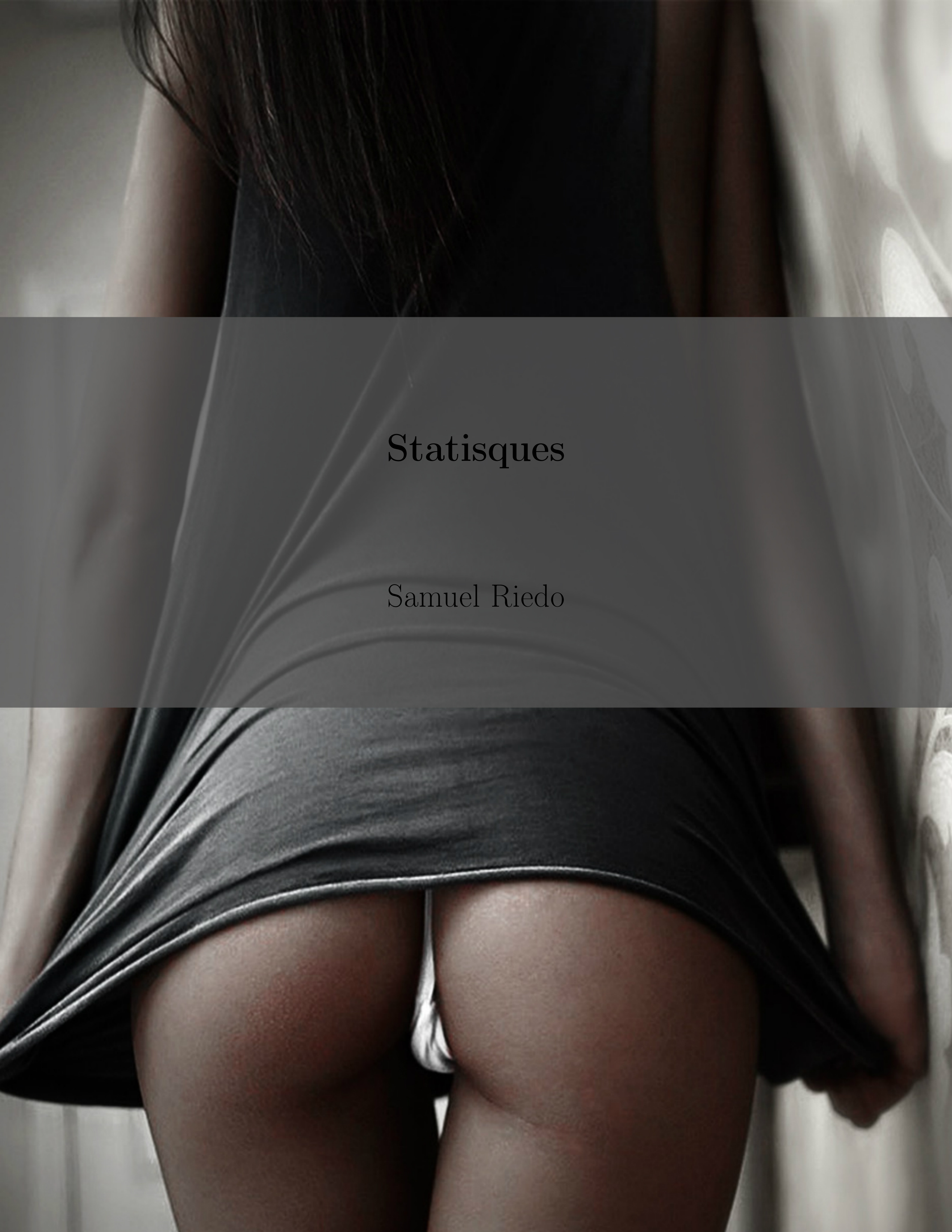


# Statisques

Samuel Riedo

# Table des matières

# 1 Statistiques descriptives

Les statistiques descriptives servent principalement à organiser et résumer les données pour faciliter leur interprétation ainsi que les futures analyses. Par exemple, la moyenne $\overline{x}$ est un nombre résumant un ensemble de données.

$$\overline{x} = \frac{x_1 + x_2 + ... + x_n}{n}$$

La variance $s^2$ indique la dispersion entre les valeurs par rapport à la moyenne :

$$s^2 = \frac{\sum_{i=1}^{n} (x_i - \overline{x})^2}{n - 1}$$

Par exemple, si un élève fait trois notes à 4, sa moyenne vaut 4. Les échantillons n'ont aucune différence avec la moyenne, la variance vaut donc 0. Si un autre élève fait 1 et 6, il a comme moyenne 3,5. La variance vaut donc :

$$s^2 = \frac{(1 - 3.5)^2 + (6 - 3.5)^2}{2 - 1} = 12.5$$

La racine de la variance, appelée écart type, correspond à l'écart moyen entre la moyenne et les valeurs. Le sample range $n$, lui, est l'écart entre la plus petite et la plus grande valeur de la statistique.

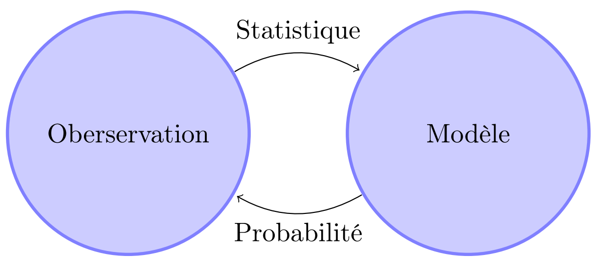
$$n = \max(x_i) - \min(x_i)$$

La médiane est une valeur divisant l'ensemble des données entre deux moitiés qui contiennent chacune le même nombre de valeurs. Si l'on divise une des deux moitiés encore une fois, les deux parties créées sont appelées 1^er^ quartile.
Le Inter Quartile Range (IQR) correspond à

$$q_3 - q_1$$

$q_3$ étant la plus petite valeur du plus grand quantile, et $q_1$ le plus grande du premier quartile.

# 2 Probabilités




# 3 Expérience aléatoire

**Problème :** Deux dés à six faces sont jetés, quelle est la probabilité d'obtenir une somme de huit ou plus ?
Première, il faut définir l'espace des échantillons.

$$S_1 = \{(1, 1), (1, 2), (1, 3), ...(6, 6)\}$$

Dans ce cas, chaque issue a la même probabilité d'apparition (1/36). Une autre modèle est :

$$S_2 = \{2, 3, ..., 12\}$$

Dans $S_2$, chaque issue n'a pas la même probabilité d'apparition.
Pour résoudre le problème, un tableau des probabilités est nécessaire.

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | | | | | |
| 2 | | | | | | × |
| 3 | | | | | × | × |
| 4 | | | | × | × | × |
| 5 | | | × | × | × | × |
| 6 | | × | × | × | × | × |

$$\text{Prob}\left(\sum \geq 8\right) = \frac{15}{36}$$

## 3.1 Règles de calcul

— Si l'absence ou la présence d'un événement $A$ n'a pas d'influence sur un autre événement $B$ (événement indépendant), alors :
$$\text{Prob}(A,\, B) = \text{Prob}(A) \cdot \text{Prob}(B)$$

— Soit $A$ un événement (A compris dans S, A C S) et $\overline{A}$ son contraire ($S \backslash A = \overline{A}$ ), alors :
$$\text{Prob}(A) + \text{Prob}(\overline{A}) = 1$$

— Principe d'adoption : Si deux événements sont incompatibles ($A \cap B = \emptyset$) alors :
$$\text{Prob}(A \cup B) = \text{Prob}(A) + \text{Prob}(B)$$

— Si une opération peut être décrite comme une suite de deux étapes avec le nombre de façons de réaliser la première étape $n_1$ et le nombre d'étapes pour la deuxième $n_2$, alors le nombre total de réaliser cette opération est :
$$n_1 \cdot n_2$$
Cette règle fonctionne également avec plus de deux étapes.

**Problème du chevalier de Méré (1654) :** Qu'est-ce qui est le plus probable ?

1. Obtenir au moins un 6 en quatre lancées de dé.
2. Obtenir au moins un double 6 en vingt-quatre lancées de deux dés.

Les chances de ne pas avoir un 6 sont :

$$\begin{aligned} P &= 1 - \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \cdot \frac{5}{6} \\ &= 0.518 \end{aligned}$$

Il y a donc 48.82% de chance d'avoir un 6. Similairement, le nombre de chance d'obtenir deux 6 en 24 vaut :

$$\begin{aligned} P &= 1 - \frac{35}{36} \cdot \frac{35}{36} \cdot \frac{35}{36} \cdot \frac{35}{36} \cdot ... \cdot \frac{35}{36} && \text{24 fois } \frac{35}{36} \\ &= 1 - \left(\frac{35}{36}\right)^{24} \\ &\approx 0.491 \end{aligned}$$

## 3.2 Probabilités conditionnelles

Soit deux événements $A$ et $B$ de l'espace des issues possibles $S$. La probabilité de $A$ étant donné que $B$ est réalisé est :

$$\text{Prob}(A|B) = \frac{\text{Prob}(A \cap B)}{\text{Prob}(B)} \qquad \text{Prob}(B) \neq 0$$

**Exemple :** Une urne avec deux boules rouges et deux boules blanches. Probabilité d'avoir deux rouges : RR

$$\begin{aligned} \text{Prob}(RR) &= \text{Prob}(*R|R*) \cdot \text{Prob}(R*) \qquad * = \text{ Don't care} \\ &= \frac{1}{3} \cdot \frac{1}{2} \\ &= \frac{1}{6} \end{aligned}$$

**Exemple :** soit des plaques en plastique ayant des défauts :
- A = Résistance au chocs (15%)
- B = Rayures (10)%
- 5% ont les deux défauts.

$$\text{Prob}(A) = 85\% \qquad \text{Prob}(B) = 80\%$$

Probabilité qu'une pièce qui soit résistance au choc si elle n'a pas de rayure ?

$$\begin{aligned} \text{Prob}(A|B) &= \frac{\text{Prob}(A \cap B)}{\text{Prob}(B)} \\ &= \frac{70}{80} \\ &= 87.5\% \end{aligned}$$

Probabilité qu'une pièce qui n'a pas de rayure soit résistance au choc ?

$$\begin{aligned} \text{Prob}(B|A) &= \frac{70}{85} \\ &= 82.35\% \end{aligned}$$

## 3.3 Loi binomiale

Des vis fabriquées par une société affectées d'un défaut avec une probabilité de 1%. Ces vis sont vendues par paquets de 10 et sont indépendantes les unes des autres. Si un paquet contient au moins deux vis défectueuses, l'entreprise le rembourse.
Quelle est la probabilité de rembourser un paquet ?

$$\begin{aligned} \text{Prob}(0) &= 0.99^{10} \cdot C_{10}^{10} \\ \text{Prob}(1) &= 0.01 \cdot 0.99^{9} \cdot C_{10}^{1} \\ \text{Prob}(2) &= 0.01^{2} \cdot 0.99^{8} \cdot C_{10}^{2} \\ \text{Prob}(k) &= 0.01^{k} \cdot 0.99^{10-k} \cdot C_{10}^{k} \end{aligned}$$

Le remboursement se fait s il y a au moins deux vis :

$$\begin{aligned} \text{Prob}(\text{remboursement}) &= 1 - \text{Prob}(0) - \text{Prob}(1) \\ &= 0.004\% \end{aligned}$$

### 3.3.1 Théorie

Lorsque l'on répète $n$ fois une expérience avec deux résultats possibles (défaut/pas de défaut, raté/réussi...) avec probabilité Prob(succès) et 1−Prob(échec) et si les résultats des expériences sont indépendants les uns des autres, on dit que le nombre de succès $x$ suit une loi binomiale de paramètres $n$ et $p$.

**Exercice :** Une urne contient $n$ boules blanches ($n \geq 5$) et 10 boules noires. On tire au hasard et simultanément 10 boules de l'urne en admettant que tous les tirages sont également probables.

1. Calculer la probabilité $P_n$ de tirer cinq boules noires et cinq seulement.
   Chacun des $C_{n+10}^{10}$ tirages possibles a la même probabilité :

$$\begin{aligned} P_n &= \frac{C_n^5 \cdot C_{10}^5}{C_{n+10}^{10}} \\ &= \frac{n! \cdot 10! \cdot 10! \cdot n!}{(n-5)! \cdot 5! \cdot 5! \cdot 5! \cdot (n+10)!} \end{aligned}$$

$x$ est une variable aléatoire comptant le nombre de succès sur $n$ expériences indépendantes, où la probabilité d'un succès vaut $p$.

$$\mathrm{Prob}(x = k) = \binom{k}{n} \cdot p^k \cdot (1-p)^{n-k} \qquad k = 0...n$$

**Exemple 4.19** Un succès est considéré si le composant fonctionne avec la probabilité $p$, $x$ désigne le nombre de composants fonctionnant dans un système à $n$ composants.

$$\begin{aligned} \mathrm{Prob}(\text{système fonctionne}) &= P(x=3) + P(x=4) + P(x=5) \\ &= \binom{5}{3} \cdot p^3 \cdot (1-p)^2 + \binom{5}{4} \cdot p^4 \cdot (1-p)^1 + p \end{aligned}$$

Avec $n$=3

$$\begin{aligned} \mathrm{Prob}(\text{système fonctionne}) &= P(x=2) + P(x=3) \\ &= \binom{3}{2} \cdot p^2 \cdot (1-p) \end{aligned}$$

Chercher

$$\binom{5}{3} \cdot p^3 \cdot (1-p)^2 + ... + p^5 > \binom{3}{2} \cdot p^2 \cdot (1-p) + p^3$$

**Exercices :**

1. Calculer la probabilité que, lorsque je lance 3x une pièce de monnaie équilibrée, il y aura
   (a) 3 fois face
   (b) 2 fois pile et une fois face
   (c) une fois pile et deux fois face
   (d) 3 fois pile

   (a) $\frac{1}{2}^3 = \frac{1}{8}$
   (b) $\frac{3}{8}$
   (c) $\frac{3}{8}$
   (d) $\frac{1}{8}$

2. Calculer la probabilité dobtenir au moins une fois un total de 7 ou plus, en jetant 3 fois une paire de dés équilibrés. Les jets sont indépendants les uns des autres.
   $X$ = le nombre de fois ou la somme est égal ou supérieur à 7.

| $E_1$ | $E_2$ | $E_3$ |
|---|---|---|
| | $P=?$ | |

   $p$ est constant.

$$P(x=1)+P(x=2)+P(x=3)=1-P(x=0)$$

| x | 0 | 1 | 2 | 3 |
|---|---|---|---|---|
| | x | ✓ | ✓ | ✓ |

   Calcule de $p$ :

| | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1 | | | | | | × |
| 2 | | | | | × | × |
| 3 | | | | × | × | × |
| 4 | | | × | × | × | × |
| 5 | | × | × | × | × | × |
| 6 | × | × | × | × | × | × |

   Il y a 21 cases cochées sur 36, donc $p=\frac{21}{36}$. Les chances d'avoir 5 7 sont :

$$\begin{aligned} P(x=0) &= 1-\binom{3}{0}\cdot\left(\frac{21}{36}\right)^0\cdot\left(\frac{15}{36}\right)^3 \\ &= 92,77\% \end{aligned}$$

3. Calculer la probabilité de deviner correctement au moins 6 des 10 réponses dun test de type juste/faux en supposant que vous répondiez au hasard.

$$C_{10}^6 \qquad\qquad P=50\%=0,5$$

$$\begin{aligned} \sum_{k=6}^{10}P(x=k) &= \sum_{k=6}^{10}\binom{10}{k}\cdot\left(\frac{1}{2}\right)^k\cdot\left(\frac{1}{2}\right)^{10-k} \\ &= 37,7\% \end{aligned}$$

4. 10% des objets produits dans une usine sont défectueux. Calculer la probabilité que dans un échantillon de 10 objets choisis au hasard, exactement 2 soient défectueux.

$$n=10 \qquad p=10\%=0,1 \qquad x=\text{ Objets défectueux}$$

$$\begin{aligned} P(x=2) &= \binom{10}{2}\cdot 0,9^8\cdot 0,1^2 \\ &= 0,1937\% \end{aligned}$$

Soit $X$ qui suit une loi binomiale avec paramètre $p$ et $n$. Lorsque $n$ est grand (>100) et $p$ petit ($n \cdot p < 5$).

$$n \cdot p = \lambda \qquad p = \frac{\lambda}{n}$$

$$p(X = k) = \frac{n! \cdot \lambda^k \cdot (1 - \frac{\lambda}{n})^n}{(n-k)! \cdot n^k \cdot (1 - \frac{\lambda}{n})^k \cdot k!} \qquad 0 \leq k \leq n$$

$$= \boxed{\frac{\lambda^k \cdot e^{-\lambda}}{k!}}$$