Some probabilistic aspects of FCA I

Richard Emilion (University of Orléans, France)

DAMOL
DATA ANALYSIS AND MODELING LAB
Palacky University, Olomouc, Czech Republic

european social fund in the czech republic
EUROPEAN UNION
MINISTRY OF EDUCATION, YOUTH AND SPORTS
OP Education for Competitiveness

INVESTMENTS IN EDUCATION DEVELOPMENT

Outline

- **Part I - Introduction**
- **Part II - Models**
- **Part III - Sampling with Markov Chains**
- **Part IV - Pointwise convergence of empirical CLs**
- **Part V - Experiments, Regression**

Part I - Introduction

Motivations, Modelling, Sampling
Basics of Probability and Statistics

I.1 Motivations

• Context $\mathcal{C} = (I, J, \mathcal{D})$ (Binary matrix case), $\mathcal{L}$ its concept lattice.

• Examples of complex and time consuming tasks: listing $\mathcal{L}$, frequent itemsets, associative rules

• Probabilistic and Statistical methods can bring a specific insight to these taks using:

1. *Modeling*
2. *Sampling, Bootstrapping*
3. *Simulation*

I.1 Motivations

• Context $\mathcal{C} = (I, J, \mathcal{D})$ (Binary matrix case), $\mathcal{L}$ its concept lattice.

• Examples of complex and time consuming tasks: listing $\mathcal{L}$, frequent itemsets, associative rules

• Probabilistic and Statistical methods can bring a specific insight to these taks using:

1. *Modeling*
2. *Sampling, Bootstrapping*
3. *Simulation*

I.1 Motivations

• Context $\mathcal{C} = (I, J, \mathcal{D})$ (Binary matrix case), $\mathcal{L}$ its concept lattice.

• Examples of complex and time consuming tasks: listing $\mathcal{L}$, frequent itemsets, associative rules

• Probabilistic and Statistical methods can bring a specific insight to these taks using:

1. *Modeling*
2. *Sampling, Bootstrapping*
3. *Simulation*

I.1 Motivations

• Context $\mathcal{C} = (I, J, \mathcal{D})$ (Binary matrix case), $\mathcal{L}$ its concept lattice.

• Examples of complex and time consuming tasks: listing $\mathcal{L}$, frequent itemsets, associative rules

• Probabilistic and Statistical methods can bring a specific insight to these taks using:

1. *Modeling*
2. *Sampling, Bootstrapping*
3. *Simulation*

I.1 Motivations

• Context $\mathcal{C} = (I, J, \mathcal{D})$ (Binary matrix case), $\mathcal{L}$ its concept lattice.

• Examples of complex and time consuming tasks: listing $\mathcal{L}$, frequent itemsets, associative rules

• Probabilistic and Statistical methods can bring a specific insight to these taks using:

1. *Modeling*

2. *Sampling, Bootstrapping*

3. *Simulation*

I.2 Modeling

- Model: Mathematical representation that can *describe/mimic* a real system
- Deterministic models, Probabilistic (Stochastic) models

- Ex. *Modeling* a real context (and $\mathcal{L}$, if possible) submitted to a random environment: customer purchases, meteorological measurements, patient diseases ...

- *Observed* measurements are considered *outcomes* of a probabilistic model.
- Statistics tasks:

- *Model Fitting*: *Estimation* of the model parameters from the observations
- Performing *Tests* and proposing *Confidence Intervals*
- *Model selection*

- Some Interest of models:

Framework for exact computations (concerning, e.g., $\mathcal{L}$) and for *prediction*
Framework for finding the true concepts and not only the empirical concepts

I.2 Modeling

- Model: Mathematical representation that can *describe/mimic* a real system
- Deterministic models, Probabilistic (Stochastic) models

• Ex. *Modeling* a real context (and $\mathcal{L}$, if possible) submitted to a random environment: customer purchases, meteorological measurements, patient diseases ...

- *Observed* measurements are considered *outcomes* of a probabilistic model.
- Statistics tasks:

\- *Model Fitting*: *Estimation* of the model parameters from the observations
\- Performing *Tests* and proposing *Confidence Intervals*
\- *Model selection*

• Some Interest of models:
Framework for exact computations (concerning, e.g., $\mathcal{L}$) and for *prediction*
Framework for finding the true concepts and not only the empirical concepts

I.2 Modeling 5 / 51

• Model: Mathematical representation that can *describe/mimic* a real system
• Deterministic models, Probabilistic (Stochastic) models

• Ex. *Modeling* a real context (and $\mathcal{L}$, if possible) submitted to a random environment: customer purchases, meteorological measurements, patient diseases ...

• *Observed* measurements are considered *outcomes* of a probabilistic model.
• Statistics tasks:
- *Model Fitting*: *Estimation* of the model parameters from the observations
- Performing *Tests* and proposing *Confidence Intervals*
- *Model selection*

• Some Interest of models:
Framework for exact computations (concerning, e.g., $\mathcal{L}$) and for *prediction*
Framework for finding the true concepts and not only the empirical concepts

I.2 Modeling

- Model: Mathematical representation that can *describe/mimic* a real system
- Deterministic models, Probabilistic (Stochastic) models

- Ex. *Modeling* a real context (and $\mathcal{L}$, if possible) submitted to a random environment: customer purchases, meteorological measurements, patient diseases ...

- *Observed* measurements are considered *outcomes* of a probabilistic model.
- Statistics tasks:

- *Model Fitting*: *Estimation* of the model parameters from the observations
- Performing *Tests* and proposing *Confidence Intervals*
- *Model selection*

- Some Interest of models:

Framework for exact computations (concerning, e.g., $\mathcal{L}$) and for *prediction*
Framework for finding the true concepts and not only the empirical concepts

I.2 Modeling

- Model: Mathematical representation that can *describe/mimic* a real system
- Deterministic models, Probabilistic (Stochastic) models

- Ex. *Modeling* a real context (and $\mathcal{L}$, if possible) submitted to a random environment: customer purchases, meteorological measurements, patient diseases ...

- *Observed* measurements are considered *outcomes* of a probabilistic model.
- Statistics tasks:

- *Model Fitting*: *Estimation* of the model parameters from the observations
- Performing *Tests* and proposing *Confidence Intervals*
- *Model selection*

- Some Interest of models:

Framework for exact computations (concerning, e.g., $\mathcal{L}$) and for *prediction*
Framework for finding the true concepts and not only the empirical concepts

I.2 Modeling

- Model: Mathematical representation that can *describe/mimic* a real system
- Deterministic models, Probabilistic (Stochastic) models

- Ex. *Modeling* a real context (and $\mathcal{L}$, if possible) submitted to a random environment: customer purchases, meteorological measurements, patient diseases ...

- *Observed* measurements are considered *outcomes* of a probabilistic model.
- Statistics tasks:

- *Model Fitting*: *Estimation* of the model parameters from the observations
- Performing *Tests* and proposing *Confidence Intervals*
- *Model selection*

- Some Interest of models:

Framework for exact computations (concerning, e.g., $\mathcal{L}$) and for *prediction*
Framework for finding the true concepts and not only the empirical concepts

I.2 Modeling

- Model: Mathematical representation that can *describe/mimic* a real system
- Deterministic models, Probabilistic (Stochastic) models

- Ex. *Modeling* a real context (and $\mathcal{L}$, if possible) submitted to a random environment: customer purchases, meteorological measurements, patient diseases ...

- *Observed* measurements are considered *outcomes* of a probabilistic model.
- Statistics tasks:

\- *Model Fitting*: *Estimation* of the model parameters from the observations
\- Performing *Tests* and proposing *Confidence Intervals*
\- *Model selection*

- Some Interest of models:

Framework for exact computations (concerning, e.g., $\mathcal{L}$) and for *prediction*
Framework for finding the true concepts and not only the empirical concepts

I.3 Sampling

• Consider a given $\mathcal{C}$ or $\mathcal{L}$ as a population: *Sample, Bootstrap* individuals from $\mathcal{C}$ or from $\mathcal{L}$

• Application : *Concept Counting* (estimating $|\mathcal{L}|$), and quickly check the feasibility of a potentially exponential time listing of all concepts

I.3 Sampling

• Consider a given $\mathcal{C}$ or $\mathcal{L}$ as a population: *Sample, Bootstrap* individuals from $\mathcal{C}$ or from $\mathcal{L}$
• Application : *Concept Counting* (estimating $|\mathcal{L}|$), and quickly check the feasibility of a potentially exponential time listing of all concepts

I.4 Measurable spaces

- Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
- Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:

$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation $(^c)$, and countable union $(\cup_n)$
Measurable space: $(\Omega, \mathcal{F})$

- Elements of $\mathcal{F}$ are called measurable sets

Examples :
- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m \times n}(\mathbb{R})$

I.4 Measurable spaces

• Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
• Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:
$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation $(^c)$, and countable union $(\cup_n)$
Measurable space: $(\Omega, \mathcal{F})$
• Elements of $\mathcal{F}$ are called measurable sets

Examples :
- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m \times n}(\mathbb{R})$

I.4 Measurable spaces

- Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
- Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:

$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation (^{c}), and countable union ($\cup_n$)

Measurable space: $(\Omega, \mathcal{F})$

- Elements of $\mathcal{F}$ are called measurable sets

Examples :

- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m \times n}(\mathbb{R})$

I.4 Measurable spaces

- Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
- Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:

$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation $(^c)$, and countable union $(\cup_n)$

Measurable space: $(\Omega, \mathcal{F})$

- Elements of $\mathcal{F}$ are called measurable sets

Examples :

- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m \times n}(\mathbb{R})$

I.4 Measurable spaces

- Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
- Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:

$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation $({}^c)$, and countable union $(\cup_n)$

Measurable space: $(\Omega, \mathcal{F})$

- Elements of $\mathcal{F}$ are called measurable sets

Examples :

- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m \times n}(\mathbb{R})$

I.4 Measurable spaces

• Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
• Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:
$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation $({}^c)$, and countable union $(\cup_n)$
Measurable space: $(\Omega, \mathcal{F})$
• Elements of $\mathcal{F}$ are called measurable sets

Examples :
- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m \times n}(\mathbb{R})$

• Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
• Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:
$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation $(^c)$, and countable union $(\cup_n)$
Measurable space: $(\Omega, \mathcal{F})$
• Elements of $\mathcal{F}$ are called measurable sets

Examples :
- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m\times n}(\mathbb{R})$

I.4 Measurable spaces

• Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
• Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:
$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation (^{c}), and countable union ($\cup_n$)
Measurable space: $(\Omega, \mathcal{F})$
• Elements of $\mathcal{F}$ are called measurable sets

Examples :
- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m \times n}(\mathbb{R})$

- Let Ω a nonvoid set and $\mathcal{P}(\Omega)$ its power set
- Let $\mathcal{F}$ be a σ-algebra on Ω, i.e:

$\mathcal{F} \subseteq \mathcal{P}(\Omega), \emptyset \in \mathcal{F}$, stable by complementation $(^c)$, and countable union $(\cup_n)$
Measurable space: $(\Omega, \mathcal{F})$

- Elements of $\mathcal{F}$ are called measurable sets

Examples :

- Ω countable (finite set, $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$) and $\mathcal{F} = \mathcal{P}(\Omega)$
- intersection of a family of σ-algebra
- Let $\mathcal{A} \subseteq \mathcal{P}(\Omega)$, the intersection of all the σ-algebras containing $\mathcal{A}$ is the smallest σ-algebra containing $\mathcal{A}$, it is called the σ-algebra generated by $\mathcal{A}$ and is denoted by $\sigma(\mathcal{A})$
- $\Omega = \mathcal{R}$, $\mathcal{F} = \sigma(\mathcal{I}), \mathcal{I}$ denoting the set of all intervals
- $\Omega = \Omega_1 \times \Omega_2$, $\mathcal{F} = \sigma(\{F_1 \times F_2, F_1 \in \mathcal{F}_1, F_2 \in \mathcal{F}_2\}$ is denoted $\mathcal{F}_1 \otimes \mathcal{F}_2$
- $\Omega = \mathbb{R}^2, \mathbb{R}^d, \mathcal{M}_{m\times n}(\mathbb{R})$

• A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that
$\mathbb{P}(\Omega) = 1, \mathbb{P}(\cup A_n) = \sum_n \mathbb{P}(A_n)$ for pairwise disjoint A_n
Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$

• *Examples* Ω finite $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$. $\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}, \mathbb{P}(\omega_n) = p_n \geq 0, \sum_n p_n = 1$.
$\Omega = [0,1]$, $\mathbb{P}([a,b]) = b - a, 0 \leq a \leq b \leq 1$

• A nullset N is a $N \in \mathcal{F} : \mathbb{P}(N) = 0$
• Let $prop$ be a property which an element $\omega \in \Omega$ may or may not have. We will say that $prop$ holds almost everyhere (a.e.) if $\{\omega \in \Omega : prop(\omega)$ is false$\}$ is a $nullset$
• A family of events $A_i \in \mathcal{F}, i \in I$ is independent if for any finite subset $J \subseteq I$ we have $\mathbb{P}(\cap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$
• A family $\mathcal{F}_i, i \in I$ of sub σ-algebras of $\mathcal{F}$ is independent if any family $A_i, i \in I$, with $A_i \in \mathcal{F}_i$, is independent

• A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that
$\mathbb{P}(\Omega) = 1, \mathbb{P}(\cup A_n) = \sum_n \mathbb{P}(A_n)$ for pairwise disjoint A_n
Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$

• *Examples* Ω finite $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$. $\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}, \mathbb{P}(\omega_n) = p_n \geq 0, \sum_n p_n = 1.$
$\Omega = [0,1]$, $\mathbb{P}([a,b]) = b - a, 0 \leq a \leq b \leq 1$

• A nullset N is a $N \in \mathcal{F} : \mathbb{P}(N) = 0$
• Let $prop$ be a property which an element $\omega \in \Omega$ may or may not have. We will say that $prop$ holds almost everyhere (a.e.) if $\{\omega \in \Omega : prop(\omega)$ is false$\}$ is a $nullset$
• A family of events $A_i \in \mathcal{F}, i \in I$ is independent if for any finite subset $J \subseteq I$ we have $\mathbb{P}(\cap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$
• A family $\mathcal{F}_i, i \in I$ of sub σ-algebras of $\mathcal{F}$ is independent if any family $A_i, i \in I$, with $A_i \in \mathcal{F}_i$, is independent

• A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that
$\mathbb{P}(\Omega) = 1, \mathbb{P}(\cup A_n) = \sum_n \mathbb{P}(A_n)$ for pairwise disjoint A_n
Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$

• *Examples* Ω finite $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$. $\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}, \mathbb{P}(\omega_n) = p_n \geq 0, \sum_n p_n = 1$.
$\Omega = [0,1]$, $\mathbb{P}([a,b]) = b - a, 0 \leq a \leq b \leq 1$

• A nullset N is a $N \in \mathcal{F} : \mathbb{P}(N) = 0$
• Let $prop$ be a property which an element $\omega \in \Omega$ may or may not have. We will say that $prop$ holds almost everyhere (a.e.) if $\{\omega \in \Omega : prop(\omega)$ is false$\}$ is a $nullset$
• A family of events $A_i \in \mathcal{F}, i \in I$ is independent if for any finite subset $J \subseteq I$ we have $\mathbb{P}(\cap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$
• A family $\mathcal{F}_i, i \in I$ of sub σ-algebras of $\mathcal{F}$ is independent if any family $A_i, i \in I$, with $A_i \in \mathcal{F}_i$, is independent

I.5 Probability spaces

• A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that
$\mathbb{P}(\Omega) = 1, \mathbb{P}(\cup A_n) = \sum_n \mathbb{P}(A_n)$ for pairwise disjoint A_n
Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$

• *Examples* Ω finite $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$. $\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}, \mathbb{P}(\omega_n) = p_n \geq 0, \sum_n p_n = 1$.
$\Omega = [0,1]$, $\mathbb{P}([a,b]) = b - a, 0 \leq a \leq b \leq 1$

• A nullset N is a $N \in \mathcal{F} : \mathbb{P}(N) = 0$
• Let $prop$ be a property which an element $\omega \in \Omega$ may or may not have. We will say that $prop$ holds almost everyhere (a.e.) if $\{\omega \in \Omega : prop(\omega)$ is false$\}$ is a $nullset$
• A family of events $A_i \in \mathcal{F}, i \in I$ is independent if for any finite subset $J \subseteq I$ we have $\mathbb{P}(\cap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$
• A family $\mathcal{F}_i, i \in I$ of sub σ-algebras of $\mathcal{F}$ is independent if any family $A_i, i \in I$, with $A_i \in \mathcal{F}_i$, is independent

I.5 Probability spaces

• A probability measure $\mathbb{P}$ on $(\Omega, \mathcal{F})$ is a mapping $\mathbb{P} : \mathcal{F} \longrightarrow [0,1]$ such that
$\mathbb{P}(\Omega) = 1, \mathbb{P}(\cup A_n) = \sum_n \mathbb{P}(A_n)$ for pairwise disjoint A_n
Probability space: $(\Omega, \mathcal{F}, \mathbb{P})$

• *Examples* Ω finite $\mathbb{P}(A) = \frac{|A|}{|\Omega|}$. $\Omega = \{\omega_1, \ldots, \omega_n, \ldots\}, \mathbb{P}(\omega_n) = p_n \geq 0, \sum_n p_n = 1$.
$\Omega = [0,1]$, $\mathbb{P}([a,b]) = b - a, 0 \leq a \leq b \leq 1$

• A nullset N is a $N \in \mathcal{F} : \mathbb{P}(N) = 0$
• Let $prop$ be a property which an element $\omega \in \Omega$ may or may not have. We will say that $prop$ holds almost everyhere (a.e.) if $\{\omega \in \Omega : prop(\omega)$ is false$\}$ is a $nullset$
• A family of events $A_i \in \mathcal{F}, i \in I$ is independent if for any finite subset $J \subseteq I$ we have $\mathbb{P}(\cap_{j \in J} A_j) = \prod_{j \in J} \mathbb{P}(A_j)$
• A family $\mathcal{F}_i, i \in I$ of sub σ-algebras of $\mathcal{F}$ is independent if any family $A_i, i \in I$, with $A_i \in \mathcal{F}_i$, is independent

I.6 Random variables

• Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(V, \mathcal{V})$ a measurable space. A V-valued random variable is a mapping $X : \Omega \longrightarrow V$ which is measurable, i.e. $X^{-1}(\mathcal{V}) \subseteq \mathcal{F}$

• Recall that if $B \in \mathcal{V}, X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\}$
$X^{-1}(\mathcal{V}) = \{X^{-1}(B), B \in \mathcal{V}\}$ is the smallest σ-algebra which makes X measurable
$X^{-1}(\mathcal{V})$ is denoted $\sigma(X)$: σ-algebra *generated* by X

• $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$ defines a probability measure on $\mathcal{V}$ called the (probability) *distribution* of X: shortly $X \sim \mathbb{P}_X$
Examples $V = \mathbb{N}, \mathbb{P}_X(\{n\}) = \mathbb{P}(X = n) = e^{-\theta}\frac{\theta^n}{n!}$. $X \sim Poisson(\theta), \theta > 0$
$V = \mathbb{R}, \mathbb{P}_X(B) = \mathbb{P}(X \in B) = \int_B f(x)dx$ with $f \geq 0, \int_{\mathbb{R}} f(x)dx = 1$.

• $X_i, i \in I$ are said independent (i) iff the σ-algebras $\sigma(X_i)$ are independent
• $X_i, i \in I$ are said identically distributed (i.d.) iff $\mathbb{P}_{X_i} = \mathbb{P}_{X_{i'}}$ for all $i, i' \in I$

I.6 Random variables

• Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(V, \mathcal{V})$ a measurable space. A V-valued random variable is a mapping $X : \Omega \longrightarrow V$ which is measurable, i.e. $X^{-1}(\mathcal{V}) \subseteq \mathcal{F}$

• Recall that if $B \in \mathcal{V}, X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\}$
$X^{-1}(\mathcal{V}) = \{X^{-1}(B), B \in \mathcal{V}\}$ is the smallest σ-algebra which makes X measurable
$X^{-1}(\mathcal{V})$ is denoted $\sigma(X)$: σ-algebra *generated* by X

• $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$ defines a probability measure on $\mathcal{V}$ called the (probability) *distribution* of X: shortly $X \sim \mathbb{P}_X$
Examples $V = \mathbb{N}, \mathbb{P}_X(\{n\}) = \mathbb{P}(X = n) = e^{-\theta}\frac{\theta^n}{n!}$. $X \sim Poisson(\theta), \theta > 0$
$V = \mathbb{R}, \mathbb{P}_X(B) = \mathbb{P}(X \in B) = \int_B f(x)dx$ with $f \geq 0, \int_{\mathbb{R}} f(x)dx = 1$.

• $X_i, i \in I$ are said independent (i) iff the σ-algebras $\sigma(X_i)$ are independent
• $X_i, i \in I$ are said identically distributed (i.d.) iff $\mathbb{P}_{X_i} = \mathbb{P}_{X_{i'}}$ for all $i, i' \in I$

I.6 Random variables

• Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(V, \mathcal{V})$ a measurable space. A V-valued random variable is a mapping $X : \Omega \longrightarrow V$ which is measurable, i.e. $X^{-1}(\mathcal{V}) \subseteq \mathcal{F}$

• Recall that if $B \in \mathcal{V}, X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\}$
$X^{-1}(\mathcal{V}) = \{X^{-1}(B), B \in \mathcal{V}\}$ is the smallest σ-algebra which makes X measurable
$X^{-1}(\mathcal{V})$ is denoted $\sigma(X)$: σ-algebra *generated* by X

• $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$ defines a probability measure on $\mathcal{V}$ called the (probability) *distribution* of X: shortly $X \sim \mathbb{P}_X$
Examples $V = \mathbb{N}, \mathbb{P}_X(\{n\}) = \mathbb{P}(X = n) = e^{-\theta}\frac{\theta^n}{n!}$. $X \sim Poisson(\theta), \theta > 0$
$V = \mathbb{R}, \mathbb{P}_X(B) = \mathbb{P}(X \in B) = \int_B f(x)dx$ with $f \geq 0, \int_{\mathbb{R}} f(x)dx = 1$.

• $X_i, i \in I$ are said independent (i) iff the σ-algebras $\sigma(X_i)$ are independent
• $X_i, i \in I$ are said identically distributed (i.d.) iff $\mathbb{P}_{X_i} = \mathbb{P}_{X_{i'}}$ for all $i, i' \in I$

I.6 Random variables

- Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space and $(V, \mathcal{V})$ a measurable space. A V-valued random variable is a mapping $X : \Omega \longrightarrow V$ which is measurable, i.e. $X^{-1}(\mathcal{V}) \subseteq \mathcal{F}$

- Recall that if $B \in \mathcal{V}, X^{-1}(B) = \{\omega \in \Omega : X(\omega) \in B\} = \{X \in B\}$
$X^{-1}(\mathcal{V}) = \{X^{-1}(B), B \in \mathcal{V}\}$ is the smallest σ-algebra which makes X measurable
$X^{-1}(\mathcal{V})$ is denoted $\sigma(X)$: σ-algebra *generated* by X

- $\mathbb{P}_X(B) = \mathbb{P}(X^{-1}(B))$ defines a probability measure on $\mathcal{V}$ called the (probability) *distribution* of X: shortly $X \sim \mathbb{P}_X$
Examples $V = \mathbb{N}, \mathbb{P}_X(\{n\}) = \mathbb{P}(X = n) = e^{-\theta}\frac{\theta^n}{n!}$. $X \sim Poisson(\theta), \theta > 0$
$V = \mathbb{R}, \mathbb{P}_X(B) = \mathbb{P}(X \in B) = \int_B f(x)dx$ with $f \geq 0, \int_{\mathbb{R}} f(x)dx = 1$.

- $X_i, i \in I$ are said independent (i) iff the σ-algebras $\sigma(X_i)$ are independent
- $X_i, i \in I$ are said identically distributed (i.d.) iff $\mathbb{P}_{X_i} = \mathbb{P}_{X_{i'}}$ for all $i, i' \in I$

I.7 Statistics, Computer simulation

• *Outcome* of X : $X(\omega)$ for some $\omega \in \Omega$. Can be any object : number, function, fuzzy set,

• A sample of X : r.v.s $X_1, \ldots X_n : X_i \overset{i.i.d.}{\sim} \mathbb{P}_X$
Observed sample : outcome $x_1, \ldots x_n$ of a sample, i.e. $x_i = X_i(\omega)$

• A *statistic* is a measurable function $T(X_1, \ldots X_n)$ of a sample: it is a r.v.
Examples $\bar{X} = \frac{X_1 + \ldots + X_n}{n}$, $S^2 = \frac{(X_1 - \bar{X})^2 + \ldots + (X_n - \bar{X})^2}{n}$, $S = \sqrt{S^2}$.
Observed statistic $T(x_1, \ldots x_n)$: it is a number.
Examples $\bar{x} = \frac{x_1 + \ldots + x_n}{n}$, $s^2 = \frac{(x_1 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}$, $s = \sqrt{s^2}$.
• LNL: For almost all (a.a.) ω, $\bar{X}(\omega) \longrightarrow \mathbb{E}(X)$ as $n \to \infty$

• Simulation: algorithm whose outputs mimic outcomes of a sample. Ex. (R)
$x = runif(100) \to x_1, \ldots, x_{100}$ with $X_i \overset{i.i.d.}{\sim} \lambda_{[0,1]}$, mean$(x) \simeq 0.5$ *for a.a.* simulations
$p = c(0.3, 0.5, 0.2), x = sample(1:3, 100, prob = p)$: sample from a discrete r.v.
$y = rnorm(100) \to y_1, \ldots, y_{100}$ with $Y_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, mean$(y) \simeq 0$ *for a.a.* simulations

• *Outcome* of X : $X(\omega)$ for some $\omega \in \Omega$. Can be any object : number, function, fuzzy set,

• A sample of X : r.v.s $X_1, \dots X_n$: $X_i \overset{i.i.d.}{\sim} \mathbb{P}_X$

Observed sample : outcome $x_1, \dots x_n$ of a sample, i.e. $x_i = X_i(\omega)$

• A *statistic* is a measurable function $T(X_1, \dots X_n)$ of a sample: it is a r.v.

Examples $\bar{X} = \frac{X_1+\dots+X_n}{n}, S^2 = \frac{(X_1-\bar{X})^2+\dots+(X_n-\bar{X})^2}{n}, \; S = \sqrt{S^2}$.

Observed statistic $T(x_1, \dots x_n)$: it is a number.

Examples $\bar{x} = \frac{x_1+\dots+x_n}{n}, s^2 = \frac{(x_1-\bar{x})^2+\dots+(x_n-\bar{x})^2}{n}, \; s = \sqrt{s^2}$.

• LNL: For almost all (a.a.) ω, $\bar{X}(\omega) \longrightarrow \mathbb{E}(X)$ as $n \to \infty$

• Simulation: algorithm whose outputs mimic outcomes of a sample. Ex. (R)

$x = runif(100) \to x_1, \dots, x_{100}$ with $X_i \overset{i.i.d.}{\sim} \lambda_{[0,1]}$, mean(x) $\simeq 0.5$ *for a.a.* simulations

$p = c(0.3, 0.5, 0.2), x = sample(1:3, 100, prob = p)$: sample from a discrete r.v.

$y = rnorm(100) \to y_1, \dots, y_{100}$ with $Y_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, mean($y$) $\simeq 0$ *for a.a.* simulations

- *Outcome* of X : $X(\omega)$ for some $\omega \in \Omega$. Can be any object : number, function, fuzzy set,
- A sample of X : r.v.s $X_1, \ldots X_n$: $X_i \overset{i.i.d.}{\sim} \mathbb{P}_X$

Observed sample : outcome $x_1, \ldots x_n$ of a sample, i.e. $x_i = X_i(\omega)$

- A *statistic* is a measurable function $T(X_1, \ldots X_n)$ of a sample: it is a r.v.

Examples $\bar{X} = \frac{X_1+\ldots+X_n}{n}, S^2 = \frac{(X_1-\bar{X})^2+\ldots+(X_n-\bar{X})^2}{n}, \ S = \sqrt{S^2}$.

Observed statistic $T(x_1, \ldots x_n)$: it is a number.

Examples $\bar{x} = \frac{x_1+\ldots+x_n}{n}, s^2 = \frac{(x_1-\bar{x})^2+\ldots+(x_n-\bar{x})^2}{n}, \ s = \sqrt{s^2}$.

- LNL: For almost all (a.a.) ω, $\bar{X}(\omega) \longrightarrow \mathbb{E}(X)$ as $n \to \infty$
- Simulation: algorithm whose outputs mimic outcomes of a sample. Ex. (R)

$x = runif(100) \to x_1, \ldots, x_{100}$ with $X_i \overset{i.i.d.}{\sim} \lambda_{[0,1]}$, mean$(x) \simeq 0.5$ *for a.a.* simulations

$p = c(0.3, 0.5, 0.2), x = sample(1:3, 100, prob = p)$: sample from a discrete r.v.

$y = rnorm(100) \to y_1, \ldots, y_{100}$ with $Y_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, mean$(y) \simeq 0$ *for a.a.* simulations

I.7 Statistics, Computer simulation

- *Outcome* of X : $X(\omega)$ for some $\omega \in \Omega$. Can be any object : number, function, fuzzy set,
- A sample of X : r.v.s $X_1, \dots X_n$: $X_i \overset{i.i.d.}{\sim} \mathbb{P}_X$

Observed sample : outcome $x_1, \dots x_n$ of a sample, i.e. $x_i = X_i(\omega)$

- A *statistic* is a measurable function $T(X_1, \dots X_n)$ of a sample: it is a r.v.

Examples $\bar{X} = \frac{X_1 + \dots + X_n}{n}$, $S^2 = \frac{(X_1 - \bar{X})^2 + \dots + (X_n - \bar{X})^2}{n}$, $S = \sqrt{S^2}$.

Observed statistic $T(x_1, \dots x_n)$: it is a number.

Examples $\bar{x} = \frac{x_1 + \dots + x_n}{n}$, $s^2 = \frac{(x_1 - \bar{x})^2 + \dots + (x_n - \bar{x})^2}{n}$, $s = \sqrt{s^2}$.

- LNL: For almost all (a.a.) ω, $\bar{X}(\omega) \longrightarrow \mathbb{E}(X)$ as $n \to \infty$

- Simulation: algorithm whose outputs mimic outcomes of a sample. Ex. (R)

$x = runif(100) \to x_1, \dots, x_{100}$ with $X_i \overset{i.i.d.}{\sim} \lambda_{[0,1]}$, mean$(x) \simeq 0.5$ *for a.a.* simulations

$p = c(0.3, 0.5, 0.2), x = sample(1:3, 100, prob = p)$: sample from a discrete r.v.

$y = rnorm(100) \to y_1, \dots, y_{100}$ with $Y_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, mean$(y) \simeq 0$ *for a.a.* simulations

I.7 Statistics, Computer simulation

- *Outcome* of X : $X(\omega)$ for some $\omega \in \Omega$. Can be any object : number, function, fuzzy set,
- A sample of X : r.v.s $X_1, \ldots X_n$: $X_i \overset{i.i.d.}{\sim} \mathbb{P}_X$

Observed sample : outcome $x_1, \ldots x_n$ of a sample, i.e. $x_i = X_i(\omega)$

- A *statistic* is a measurable function $T(X_1, \ldots X_n)$ of a sample: it is a r.v.

Examples $\bar{X} = \frac{X_1 + \ldots + X_n}{n}$, $S^2 = \frac{(X_1 - \bar{X})^2 + \ldots + (X_n - \bar{X})^2}{n}$, $S = \sqrt{S^2}$.

Observed statistic $T(x_1, \ldots x_n)$: it is a number.

Examples $\bar{x} = \frac{x_1 + \ldots + x_n}{n}$, $s^2 = \frac{(x_1 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}$, $s = \sqrt{s^2}$.

- LNL: For almost all (a.a.) ω, $\bar{X}(\omega) \longrightarrow \mathbb{E}(X)$ as $n \to \infty$

- Simulation: algorithm whose outputs mimic outcomes of a sample. Ex. (R)

$x = runif(100) \to x_1, \ldots, x_{100}$ with $X_i \overset{i.i.d.}{\sim} \lambda_{[0,1]}$, mean$(x) \simeq 0.5$ *for a.a.* simulations

$p = c(0.3, 0.5, 0.2), x = sample(1:3, 100, prob = p)$: sample from a discrete r.v.

$y = rnorm(100) \to y_1, \ldots, y_{100}$ with $Y_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, mean$(y) \simeq 0$ *for a.a.* simulations

- *Outcome* of X : $X(\omega)$ for some $\omega \in \Omega$. Can be any object : number, function, fuzzy set,
- A sample of X : r.v.s $X_1, \ldots X_n : X_i \overset{i.i.d.}{\sim} \mathbb{P}_X$

Observed sample : outcome $x_1, \ldots x_n$ of a sample, i.e. $x_i = X_i(\omega)$

- A *statistic* is a measurable function $T(X_1, \ldots X_n)$ of a sample: it is a r.v.

Examples $\bar{X} = \frac{X_1 + \ldots + X_n}{n}, S^2 = \frac{(X_1 - \bar{X})^2 + \ldots + (X_n - \bar{X})^2}{n}, \; S = \sqrt{S^2}$.

Observed statistic $T(x_1, \ldots x_n)$: it is a number.

Examples $\bar{x} = \frac{x_1 + \ldots + x_n}{n}, s^2 = \frac{(x_1 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}, \; s = \sqrt{s^2}$.

- LNL: For almost all (a.a.) ω, $\bar{X}(\omega) \longrightarrow \mathbb{E}(X)$ as $n \to \infty$
- Simulation: algorithm whose outputs mimic outcomes of a sample. Ex. (R)

$x = runif(100) \to x_1, \ldots, x_{100}$ with $X_i \overset{i.i.d.}{\sim} \lambda_{[0,1]}$, mean$(x) \simeq 0.5$ *for a.a.* simulations

$p = c(0.3, 0.5, 0.2), x = sample(1:3, 100, prob = p)$: sample from a discrete r.v.

$y = rnorm(100) \to y_1, \ldots, y_{100}$ with $Y_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, mean$(y) \simeq 0$ *for a.a.* simulations

I.7 Statistics, Computer simulation 10 / 51

- *Outcome* of X : $X(\omega)$ for some $\omega \in \Omega$. Can be any object : number, function, fuzzy set,
- A sample of X : r.v.s $X_1, \ldots X_n$: $X_i \overset{i.i.d.}{\sim} \mathbb{P}_X$

Observed sample : outcome $x_1, \ldots x_n$ of a sample, i.e. $x_i = X_i(\omega)$

- A *statistic* is a measurable function $T(X_1, \ldots X_n)$ of a sample: it is a r.v.

Examples $\bar{X} = \frac{X_1 + \ldots + X_n}{n}, S^2 = \frac{(X_1 - \bar{X})^2 + \ldots + (X_n - \bar{X})^2}{n}$, $S = \sqrt{S^2}$.

Observed statistic $T(x_1, \ldots x_n)$: it is a number.

Examples $\bar{x} = \frac{x_1 + \ldots + x_n}{n}, s^2 = \frac{(x_1 - \bar{x})^2 + \ldots + (x_n - \bar{x})^2}{n}$, $s = \sqrt{s^2}$.

- LNL: For almost all (a.a.) ω, $\bar{X}(\omega) \longrightarrow \mathbb{E}(X)$ as $n \to \infty$

- Simulation: algorithm whose outputs mimic outcomes of a sample. Ex. (R)

$x = runif(100) \to x_1, \ldots, x_{100}$ with $X_i \overset{i.i.d.}{\sim} \lambda_{[0,1]}$, mean$(x) \simeq 0.5$ *for a.a.* simulations

$p = c(0.3, 0.5, 0.2), x = sample(1:3, 100, prob = p)$: sample from a discrete r.v.

$y = rnorm(100) \to y_1, \ldots, y_{100}$ with $Y_i \overset{i.i.d.}{\sim} \mathcal{N}(0,1)$, mean$(y) \simeq 0$ *for a.a.* simulations

II - Models of random binary contexts

Bernoulli Model
Hierarchical Bernoulli Models
Indian Buffet
Latent Block Model

II.1 Bernoulli(p) Model: Simulation

$D = m \times n$ random binary matrix

mn entries $\overset{i.i.d.}{\sim} Bern(p)$, $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$

Illustration In R software : $D =$ matrix(rbinom(m*n,1,p), m,n)

$m = 10$ rows (customers, objects), $I = 1, \dots, m$

$n = 5$ columns (items, attributes) $J = 1, \dots, n$

$p = 0.4$: probability that an entry be equal to 1

Itemset $\{1, 4\}$ may be closed or not closed, depending on the outcome D.

II.1 Bernoulli(p) Model: Simulation

$D = m \times n$ random binary matrix

mn entries $\overset{i.i.d.}{\sim} Bern(p)$, $\mathbb{P}(X = 1) = p, \mathbb{P}(X = 0) = 1 - p$

Illustration In R software : $D =$ matrix(rbinom(m*n,1,p), m,n)

$m = 10$ rows (customers, objects), $I = 1, \ldots, m$

$n = 5$ columns (items, attributes) $J = 1, \ldots, n$

$p = 0.4$: probability that an entry be equal to 1

Itemset $\{1, 4\}$ may be closed or not closed, depending on the outcome D.

II.2 Random Concepts

p_j probability that any entry of column j be equal to 1

The entries of the matrix $\mathcal{D}$ are *independent* r.v.s.

O a subset of objects, A a subset of attributes (itemset)

Probability that the rectangle $O \times A$ be a concept ?

The rectangle $O \times A$ is a concept (maximal rectangle of ones) iff

1. $O \times A$ is filled of ones

and

2. each row of the rectangle $(I - O) \times A$ contains at least one zero

and

3. each column of the rectangle $O \times (J - A)$ contains at least one zero

p_j probability that any entry of column j be equal to 1

The entries of the matrix $\mathcal{D}$ are *independent* r.v.s.

O a subset of objects, A a subset of attributes (itemset)

Probability that the rectangle $O \times A$ be a concept ?
The rectangle $O \times A$ is a concept (maximal rectangle of ones) iff

1. $O \times A$ is filled of ones

and

2. each row of the rectangle $(I - O) \times A$ contains at least one zero

and

3. each column of the rectangle $O \times (J - A)$ contains at least one zero

p_j probability that any entry of column j be equal to 1

The entries of the matrix $\mathcal{D}$ are *independent* r.v.s.

O a subset of objects, A a subset of attributes (itemset)

Probability that the rectangle $O \times A$ be a concept ?
The rectangle $O \times A$ is a concept (maximal rectangle of ones) iff

1. $O \times A$ is filled of ones

and
2. each row of the rectangle $(I - O) \times A$ contains at least one zero

and
3. each column of the rectangle $O \times (J - A)$ contains at least one zero

II.2 Random Concepts

p_j probability that any entry of column j be equal to 1

The entries of the matrix $\mathcal{D}$ are *independent* r.v.s.

O a subset of objects, A a subset of attributes (itemset)

Probability that the rectangle $O \times A$ be a concept ?
The rectangle $O \times A$ is a concept (maximal rectangle of ones) iff

1. $O \times A$ is filled of ones

and
2. each row of the rectangle $(I - O) \times A$ contains at least one zero

and
3. each column of the rectangle $O \times (J - A)$ contains at least one zero

II.3 Computation in the Bernoulli model case

Let $p_A := \Pi_{j \in A} p_j$.

1.

One row of $O \times A$ is filled with ones with probability (w.p.): p_A

$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.

One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$

each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.

Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$

each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at

Proposition 1

$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

II.3 Computation in the Bernoulli model case

Let $p_A := \Pi_{j \in A} p_j$.

1.
One row of $O \times A$ is filled with ones with probability (w.p.): p_A
$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.
One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$
each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.
Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$
each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at
Proposition 1
$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

II.3 Computation in the Bernoulli model case

Let $p_A := \Pi_{j \in A} p_j$.

1.
One row of $O \times A$ is filled with ones with probability (w.p.): p_A
$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.
One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$
each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.
Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$
each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at
Proposition 1
$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

II.3 Computation in the Bernoulli model case

Let $p_A := \Pi_{j \in A} p_j$.

1.
One row of $O \times A$ is filled with ones with probability (w.p.): p_A
$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.
One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$
each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.
Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$
each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at
Proposition 1
$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

Let $p_A := \Pi_{j \in A} p_j$.

1.
One row of $O \times A$ is filled with ones with probability (w.p.): p_A
$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.
One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$
each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.
Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$
each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at
Proposition 1
$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

Let $p_A := \Pi_{j \in A} p_j$.

1.
One row of $O \times A$ is filled with ones with probability (w.p.): p_A
$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.
One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$
each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.
Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$
each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at
Proposition 1
$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

Let $p_A := \Pi_{j \in A} p_j$.

1.

One row of $O \times A$ is filled with ones with probability (w.p.): p_A

$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.

One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$

each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.

Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$

each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at

Proposition 1

$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

Let $p_A := \Pi_{j \in A} p_j$.

1.
One row of $O \times A$ is filled with ones with probability (w.p.): p_A
$O \times A$ is filled of ones w.p. $p_A^{|O|}$

2.
One row of $(I - O) \times A)$ contains at least one zero w.p. $1 - p_A$
each row of $(I - O) \times A)$ contains at least one zero w.p. $(1 - p_A)^{m-|O|}$

3.
Column j of $O \times (J - A)$ contains at least one zero w.p.: $1 - p_j^{|O|}$
each column of $O \times (J - A)$ contains at least one zero w.p.: $\Pi_{j \notin A}(1 - p_j^{|O|})$

Due to independency we arrive at
Proposition 1
$O \times A$ is a maximal rectangle w.p. $p_A^{|O|}(1 - p_A)^{m-|O|}\Pi_{j \notin A}(1 - p_j^{|O|})$

II.4 Probability of A be closed, in the Bernoulli model case 15 / 51

• Given A, the preceding proposition shows that the probability only depends on the size $|O|$ of O

• As Prob(A is k-closed) $= \sum_{O \in \mathcal{P}(I)}$Prob($O \times A$ is a concept)

and there are $\binom{m}{k}$ subsets O such that $|O| = k$ we arrive at the

Proposition 2
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p_A^k (1-p_A)^{m-k} \Pi_{j \notin A}(1-p_j^k)$

• If $p_j = p$ does not depend on j, we have $p_A = p^{|A|}$ and
Proposition 3
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}$

II.4 Probability of A be closed, in the Bernoulli model case 15 / 51

- Given A, the preceding proposition shows that the probability only depends on the size $|O|$ of O
- As Prob(A is k-closed) $= \sum_{O \in \mathcal{P}(I)}$Prob($O \times A$ is a concept)

and there are $\binom{m}{k}$ subsets O such that $|O| = k$ we arrive at the

Proposition 2

Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p_A^k (1-p_A)^{m-k} \Pi_{j \notin A}(1-p_j^k)$

- If $p_j = p$ does not depend on j, we have $p_A = p^{|A|}$ and

Proposition 3

Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}$

II.4 Probability of A be closed, in the Bernoulli model case 15 / 51

• Given A, the preceding proposition shows that the probability only depends on the size $|O|$ of O

• As Prob(A is k-closed) $= \sum_{O\in\mathcal{P}(I)}$Prob($O \times A$ is a concept)

and there are $\binom{m}{k}$ subsets O such that $|O| = k$ we arrive at the

Proposition 2

Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p_A^k (1-p_A)^{m-k} \Pi_{j\notin A}(1-p_j^k)$

• If $p_j = p$ does not depend on j, we have $p_A = p^{|A|}$ and

Proposition 3

Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}$

II.4 Probability of A be closed, in the Bernoulli model case 15 / 51

• Given A, the preceding proposition shows that the probability only depends on the size $|O|$ of O

• As Prob(A is k-closed) $= \sum_{O \in \mathcal{P}(I)}$Prob($O \times A$ is a concept)

and there are $\binom{m}{k}$ subsets O such that $|O| = k$ we arrive at the

Proposition 2
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p_A^k (1-p_A)^{m-k} \Pi_{j \notin A}(1-p_j^k)$

• If $p_j = p$ does not depend on j, we have $p_A = p^{|A|}$ and
Proposition 3
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}$

II.4 Probability of A be closed, in the Bernoulli model case 15 / 51

• Given A, the preceding proposition shows that the probability only depends on the size $|O|$ of O

• As Prob(A is k-closed) $= \sum_{O \in \mathcal{P}(I)}$Prob($O \times A$ is a concept)

and there are $\binom{m}{k}$ subsets O such that $|O| = k$ we arrive at the

Proposition 2
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p_A^k (1 - p_A)^{m-k} \Pi_{j \notin A} (1 - p_j^k)$

• If $p_j = p$ does not depend on j, we have $p_A = p^{|A|}$ and
Proposition 3
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p^{k|A|} (1 - p^{|A|})^{m-k} (1 - p^k)^{n-|A|}$

II.4 Probability of A be closed, in the Bernoulli model case 15 / 51

• Given A, the preceding proposition shows that the probability only depends on the size $|O|$ of O

• As Prob(A is k-closed) $= \sum_{O\in\mathcal{P}(I)}$Prob($O \times A$ is a concept)

and there are $\binom{m}{k}$ subsets O such that $|O| = k$ we arrive at the

Proposition 2
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p_A^k (1-p_A)^{m-k} \Pi_{j\notin A}(1-p_j^k)$

• If $p_j = p$ does not depend on j, we have $p_A = p^{|A|}$ and
Proposition 3
Prob(A is k-closed) $= \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}$

II.5 Expectation of $|\mathcal{L}|$ in the Bernoulli model case

- Since the number of concepts is equal to the number of k-closed itemsets, we have

$$|L| = \sum_{A \in \mathcal{P}(J)} 1_{A \text{ is } k\text{-closed}}$$

- Taking expectation we get

$$\begin{aligned}\mathbb{E}(|L|) &= \sum_{A \in \mathcal{P}(J)} prob(A \text{ is } k\text{-closed}) \\ &= \sum_{A \in \mathcal{P}(J)} \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}\end{aligned}$$

and grouping the subsets A with same cardinality we get

Theorem 1

$$\mathbb{E}(|L|) = \sum_{l=0}^{n} \binom{n}{l} \sum_{k=0}^{m} \binom{m}{k} p^{kl}(1-p^l)^{m-k}(1-p^k)^{n-l}$$

- Since the number of concepts is equal to the number of k-closed itemsets, we have

$$|L| = \sum_{A \in \mathcal{P}(J)} 1_{A \text{ is } k\text{-closed}}$$

- Taking expectation we get

$$\begin{aligned}\mathbb{E}(|L|) &= \sum_{A \in \mathcal{P}(J)} prob(A \text{ is } k\text{-closed}) \\ &= \sum_{A \in \mathcal{P}(J)} \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}\end{aligned}$$

and grouping the subsets A with same cardinality we get

Theorem 1

$$\mathbb{E}(|L|) = \sum_{l=0}^{n} \binom{n}{l} \sum_{k=0}^{m} \binom{m}{k} p^{kl}(1-p^l)^{m-k}(1-p^k)^{n-l}$$

- Since the number of concepts is equal to the number of k-closed itemsets, we have

$$|L| = \sum_{A \in \mathcal{P}(J)} 1_{A \text{ is } k\text{-closed}}$$

- Taking expectation we get

$$\begin{aligned}\mathbb{E}(|L|) &= \sum_{A \in \mathcal{P}(J)} prob(A \text{ is } k\text{-closed}) \\ &= \sum_{A \in \mathcal{P}(J)} \sum_{k=0}^{m} \binom{m}{k} p^{k|A|}(1-p^{|A|})^{m-k}(1-p^k)^{n-|A|}\end{aligned}$$

and grouping the subsets A with same cardinality we get

Theorem 1

$$\mathbb{E}(|L|) = \sum_{l=0}^{n} \binom{n}{l} \sum_{k=0}^{m} \binom{m}{k} p^{kl}(1-p^l)^{m-k}(1-p^k)^{n-l}$$

II.6 Variance of $|\mathcal{L}|$ in the Bernoulli model case

- Computation of Prob(A and B be closed), A, $B \in \mathcal{P}(J)$

Instead of having just 3 cases, namely $O \times A$, $I - O \times A$, $O \times J - A$, it appears 16 cases. Some formulas in (Emilion-Lévy can be simplified).

- Taking expectation yields $\mathbb{E}(|L|^2)$ and therefore $var(|L|) = \mathbb{E}(|L|^2) - (\mathbb{E}(|L|))^2$

- Computation of Prob(A and B be closed), A, $B \in \mathcal{P}(J)$

Instead of having just 3 cases, namely $O \times A$, $I - O \times A$, $O \times J - A$, it appears 16 cases. Some formulas in (Emilion-Lévy can be simplified).

- Taking expectation yields $\mathbb{E}(|L|^2)$ and therefore $var(|L|) = \mathbb{E}(|L|^2) - (\mathbb{E}(|L|))^2$

II.6 Variance of $|\mathcal{L}|$ in the Bernoulli model case

- Computation of Prob(A and B be closed), A, $B \in \mathcal{P}(J)$

Instead of having just 3 cases, namely $O \times A$, $I - O \times A$, $O \times J - A$, it appears 16 cases. Some formulas in (Emilion-Lévy can be simplified).

- Taking expectation yields $\mathbb{E}(|L|^2)$ and therefore $var(|L|) = \mathbb{E}(|L|^2) - (\mathbb{E}(|L|))^2$