

# Fundamentos Matemáticos del Machine Learning

Manuel Gijón Agudo

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Probabilidad</b>	<b>3</b>
2.1. Introducción . . . . .	3
2.2. Conceptos básicos . . . . .	3
2.3. Distribuciones discretas . . . . .	3
2.3.1. Bernulli, $B(1, p)$ . . . . .	3
2.3.2. Binomial, $B(n, p)$ . . . . .	3
2.3.3. Binomial Negativa, $BN(r, p)$ . . . . .	3
2.3.4. Multinomial . . . . .	3
2.3.5. Chi Cuadrado de Pearson, $\chi^2_n$ . . . . .	4
2.3.6. T de Student, $t_n$ . . . . .	4
2.3.7. F de Fisher-Snedecor, $F_{n_1, n_2}$ . . . . .	4
2.4. Teoremas y resultados . . . . .	4
<b>3. Grafos</b>	<b>5</b>
3.1. Introducción . . . . .	5
3.2. Conceptos básicos . . . . .	5
<b>4. Word2Vect</b>	<b>6</b>
4.1. Introducción . . . . .	6
4.2. The Skip-Gram model . . . . .	6
4.3. The Continuous Bag-of-Words Models (CBOW) . . . . .	6
<b>Referencias</b>	<b>7</b>

## 1. Introducción

## 2. Probabilidad

### 2.1. Introducción

### 2.2. Conceptos básicos

### 2.3. Distribuciones discretas

#### 2.3.1. Bernulli, $B(1, p)$

#### 2.3.2. Binomial, $B(n, p)$

#### 2.3.3. Binomial Negativa, $BN(r, p)$

#### 2.3.4. Multinomial

La **distribución multinomial** una generalización de la distribución binomial.

La distribución binomial es la probabilidad de un número de éxitos en  $N$  sucesos de Bernoulli independientes, con la misma probabilidad de éxito en cada suceso. En una distribución multinomial, el análogo a la distribución de Bernoulli es la distribución categórica, donde cada suceso concluye en únicamente un resultado de un número finito  $K$  de los posibles, con probabilidades  $p_1, p_2, \dots, p_k$  (tales que  $p_i \geq 0$  para  $i \in [0, k]$  y  $\sum_{i=1}^k p_i = 1$ ); y con  $n$  sucesos independientes.

Sea la variable aleatoria  $X_i$ , que indica el número de veces que se ha dado el resultado  $i$  entre los  $n$  sucesos. El vector  $X = (X_1, \dots, X_k)$  sigue una distribución multinomial con parámetros  $n$  y  $p$ , donde  $p = (p_1, \dots, p_k)$ .

■ Parámetros:

- $n \in \mathbb{N}$ : número de pruebas.
- $p_1, \dots, p_k$ : probabilidad de un suceso concreto, tales que  $\sum p_i = 1$ .

■ Dominio:  $X_i \in \{0, \dots, n\}$  tales que  $\sum X_i = n$ .

■ Función de densidad:

$$\frac{n!}{x_1! \dots x_k!} p_1^{x_1} \dots p_k^{x_k}$$

■ Media:  $\mathbb{E}(X_i) = np_i$

■ Varianza:  $\text{Var}(X_i) = np_i(1 - p_i)$

■ Covarianza:  $\text{Cov}(X_i, X_j) = -np_i p_j$ , ( $i \neq j$ )

■ Función generadora de momentos:

$$\left( \sum_{i=1}^k p_i e^{t_i} \right)^n$$

**2.3.5. Chi Cuadrado de Pearson,  $\chi_n^2$**

**2.3.6. T de Student,  $t_n$**

**2.3.7. F de Fisher-Snedecor,  $F_{n_1, n_2}$**

**2.4. Teoremas y resultados**

## **3. Grafos**

### **3.1. Introducción**

### **3.2. Conceptos básicos**

## **4. Word2Vect**

### **4.1. Introducción**

### **4.2. The Skip-Gram model**

### **4.3. The Continuous Bag-of-Words Models (CBOW)**

## Referencias

- [w2v] Yoav Goldberg and Omer Levy “word2vec Explained: Deriving Mikolov et al.’s Negative-Sampling Word-Embedding Method” **31** (February 14, 2014)
-