

Introducción al análisis cuantitativo de datos lingüísticos

Bloque 3.1: Probabilidad condicional

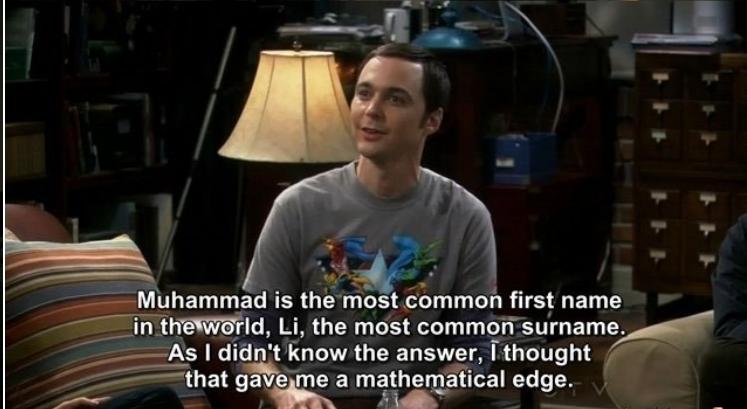
Ezequiel Koile (MPI-SSH)
Carolina Gattai (IFIBA – CONICET)

Probabilidad conjunta (más de una variable)

- $\Pr(\text{llueve}) = 0.3$
- $\Pr(\text{es martes}) = 1/7 = 0.1428\dots$
- $\Pr(\text{llueve Y es martes}) = \Pr(\text{llueve}) \times \Pr(\text{es martes})$
 $= 0.3 \times 0.1428\dots$
 $= 0.04285\dots$

¿Siempre podemos hacer esto?

Este razonamiento funciona siempre, ¿no?



Ejemplo



- ▶ Tiramos un dado
- ▶ ¿Cuál es la probabilidad de obtener un número impar?
- ▶ ¿Cuál es la probabilidad de obtener un número “bajo” (≤ 3)?
- ▶ ¿Cuál es la probabilidad de obtener un número impar, *dado que* sabemos que es bajo?
- ▶ ¿Cuál es la probabilidad de obtener un número bajo, *dado que* sabemos que es impar?
- ▶ ¿Qué pasa si redefinimos “bajo” como un número ≤ 2 ?

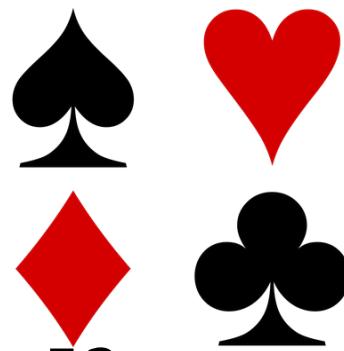
Probabilidad conjunta (más de una variable)

- ▶ $\Pr(A \text{ y } B) = \Pr(A | B) * \Pr(B)$ ←
- ▶ $\Pr(A \text{ y } B) = \Pr(B | A) * \Pr(A)$
- ▶ En nuestro caso:
 - $\Pr(\text{impar Y bajo}) = \Pr(\text{impar|bajo}) * \Pr(\text{bajo})$
 - $\Pr(\text{impar Y bajo}) = \Pr(\text{bajo|impar}) * \Pr(\text{impar})$
- ▶ Si
 - $\Pr(A|B) = \Pr(A)$
- ▶ Entonces, también se cumple que
 - $\Pr(B|A) = \Pr(B)$
- ▶ Por lo tanto
 - $\Pr(A \text{ y } B) = \Pr(A) * \Pr(B)$

En este caso, los eventos son **independientes**

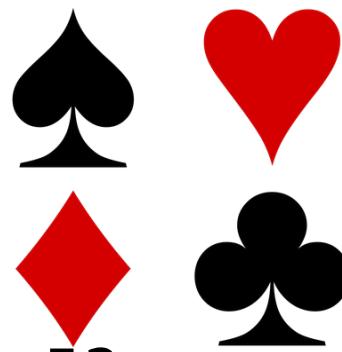
Probabilidad
condicional

Ejemplo 2



- ▶ Saco una carta de poker al azar (mazo de 52 cartas, 4 palos)
- ▶ ¿Cuál es la probabilidad de sacar un as?
- ▶ ¿Cuál es la probabilidad de sacar un diamante?
- ▶ ¿Cuál es la probabilidad de sacar un as, *dado que* sacamos un diamante?
- ▶ ¿Cuál es la probabilidad de sacar un diamante, *dado que* sacamos un as?
- ▶ Esto significa que ambos eventos son *independientes*

Ejemplo 2



- ▶ Saco una carta de poker al azar (mazo de 52 cartas, 4 palos)
- ▶ ¿Cuál es la probabilidad de sacar un as?
- ▶ ¿Cuál es la probabilidad de sacar un diamante?
- ▶ ¿Cuál es la probabilidad de sacar un as, *dado que* sacamos un diamante?
- ▶ ¿Cuál es la probabilidad de sacar un diamante, *dado que* sacamos un as?
- ▶ Esto significa que ambas probabilidades son *independientes*
- ▶ ¿Y si agregamos un comodín?



Probabilidad conjunta (más de una variable)

- ▶ Eventos no independientes: probabilidad condicional
- ▶ Ejemplo lingüístico
 - $\Pr(\text{word order} = \text{VO}) = 0.506$ (WALS)
 - $\Pr(\text{adposition} = \text{preposition}) = 0.478$ (WALS)
 - $\Pr(\text{VO AND preposition}) = ?$
 - $\Pr(\text{VO}) * \Pr(\text{preposition}) = 0.242$
 - $\Pr(\text{VO AND preposition}) = 0.463$ (WALS)

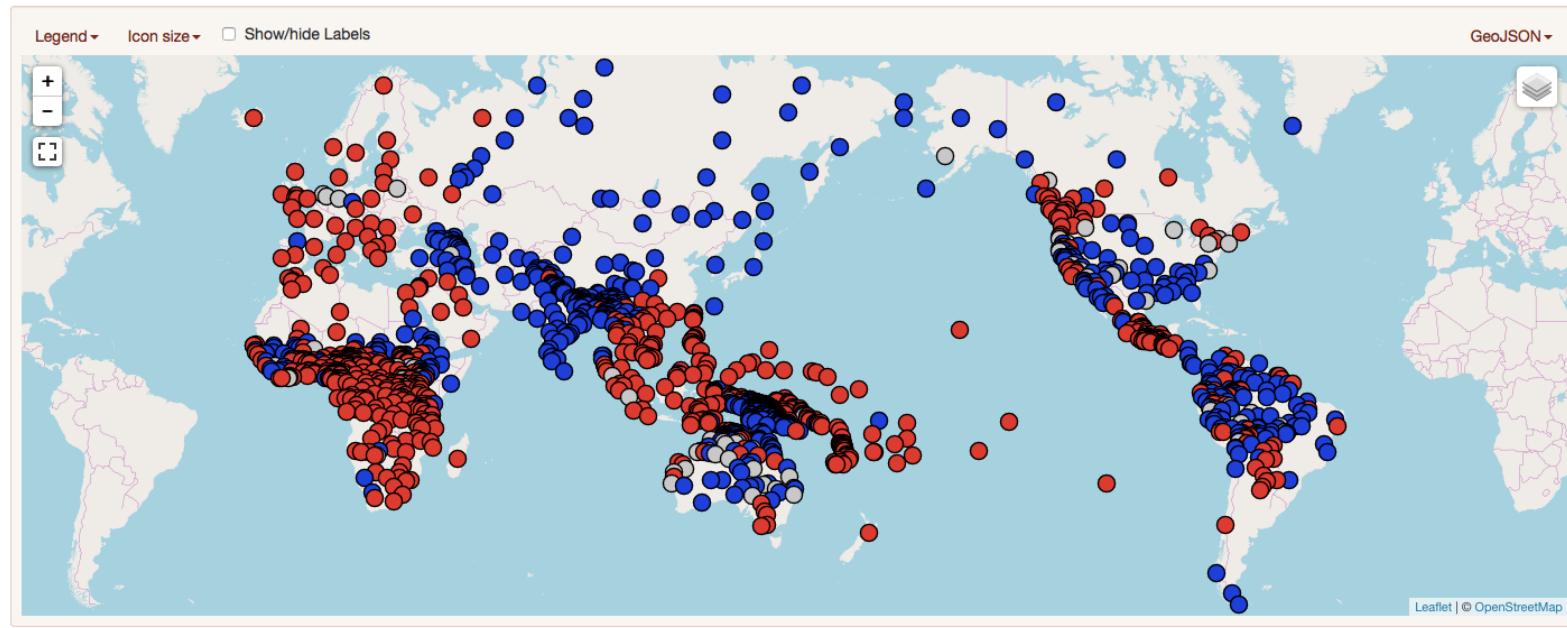
Probabilidad conjunta (más de una variable)

Feature 83A: Order of Object and Verb

This feature is described in the text of chapter 83 [Order of Object and Verb](#) by Matthew S. Dryer [cite](#)

You may combine this feature with another one. Start typing the feature name or number in the field below.

Values		
● OV	713	
● VO	705	
○ No dominant order	101	



Probabilidad conjunta (más de una variable)

Feature 85A: Order of Adposition and Noun Phrase

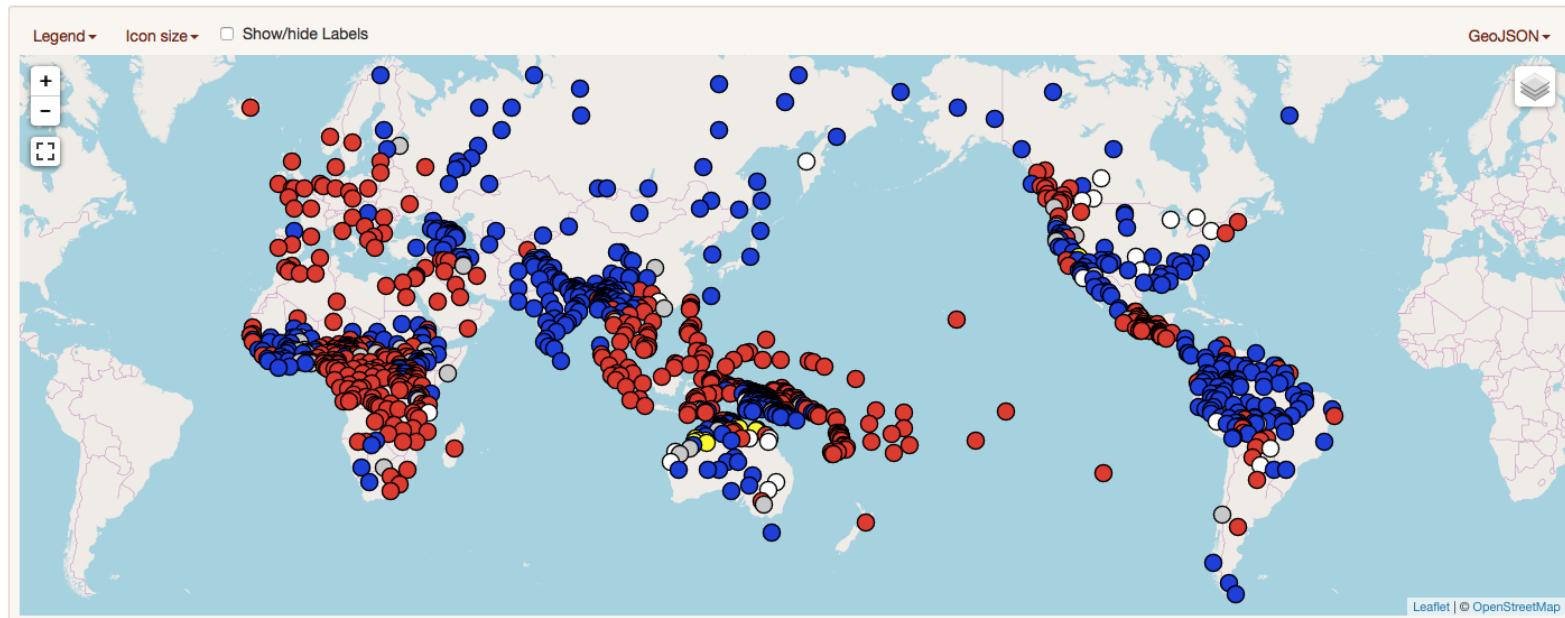


This feature is described in the text of chapter 85 [Order of Adposition and Noun Phrase](#) by Matthew S. Dryer [cite](#)

You may combine this feature with another one. Start typing the feature name or number in the field below.

Values

●	Postpositions	576
●	Prepositions	511
●	Inpositions	8
●	No dominant order	58
○	No adpositions	30



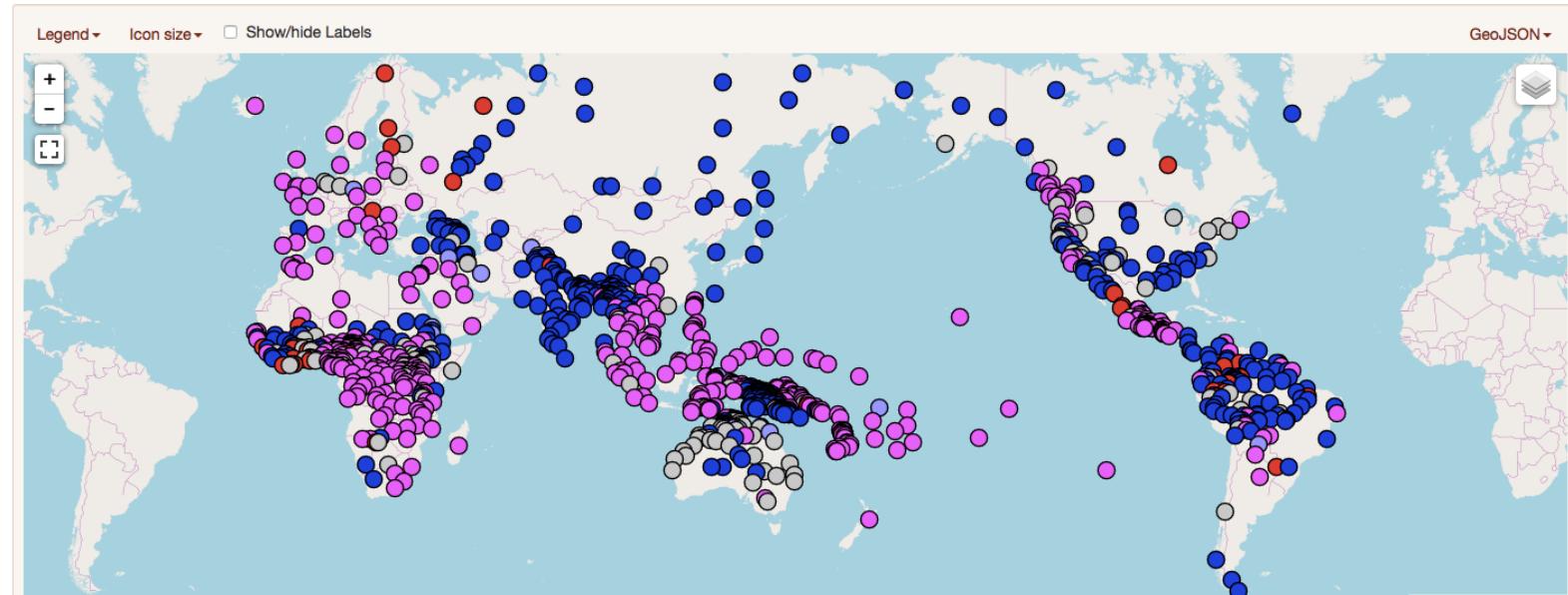
Probabilidad conjunta (más de una variable)

This feature is described in the text of chapter 95 Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase by Matthew S. Dryer [cite](#)

You may combine this feature with another one. Start typing the feature name or number in the field below.

95A: Relationship between the Order of Object and Verb and the Order of Adposition and Noun Phrase Submit

Category	Count
OV and Postpositions	472
OV and Prepositions	14
VO and Postpositions	42
VO and Prepositions	456
Other	158



Joint probability (multiple variables)

- ▶ $\Pr(A \text{ y } B) = \Pr(A | B) * \Pr(B)$
- ▶ $\Pr(A \text{ y } B) = \Pr(B | A) * \Pr(A)$

●	OV and Postpositions	472
○	OV and Prepositions	14
●	VO and Postpositions	42
●	VO and Prepositions	456
○	Other	158

- ▶ En nuestro caso
 - $\Pr(\text{VO}) = 0.506$ (WALS)
 - $\Pr(\text{preposition}) = 0.429$ (WALS)
 - $\Pr(\text{VO} | \text{preposition}) = 0.478$ (WALS)
 - $\Pr(\text{preposition} | \text{VO}) = 0.931$ (WALS)
 - $\Pr(\text{VO AND preposition}) = 0.463$ (WALS) **Check it!**

¿Y si queremos calcular
 $\Pr(\text{VO} | \text{preposition})$?

Teorema de Bayes

$$P(A | B) = \frac{P(B | A) P(A)}{P(B)}$$