Conditional Distributions of Discrete Random Variables

Probability and Statistics for Data Science

Carlos Fernandez-Granda

NYU COURANT INSTITUTE OF MATHEMATICAL SCIENCES

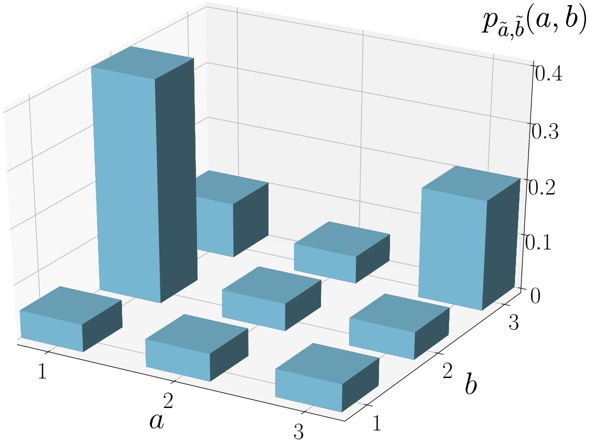
NYU DATA SCIENCE

These slides are based on the book Probability and Statistics for Data Science by Carlos Fernandez-Granda, available for purchase here. A free preprint, videos, code, slides and solutions to exercises are available at https://www.ps4ds.net

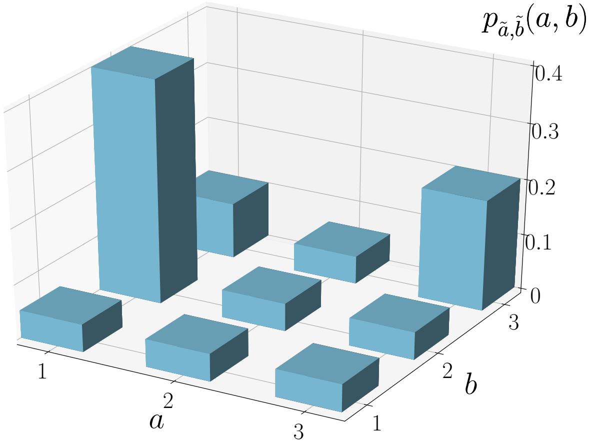
Motivation

How do we update a model if the value of some variables are revealed?

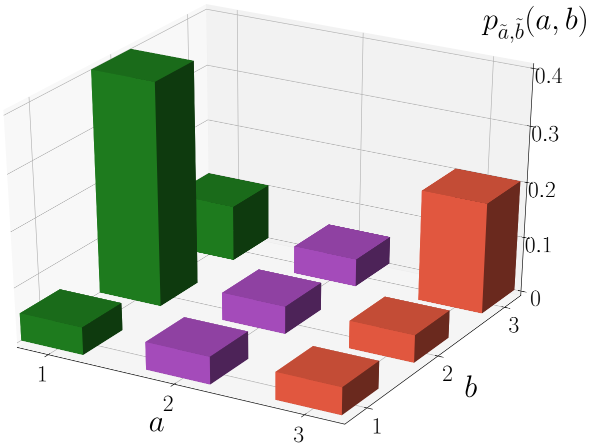
Joint pmf



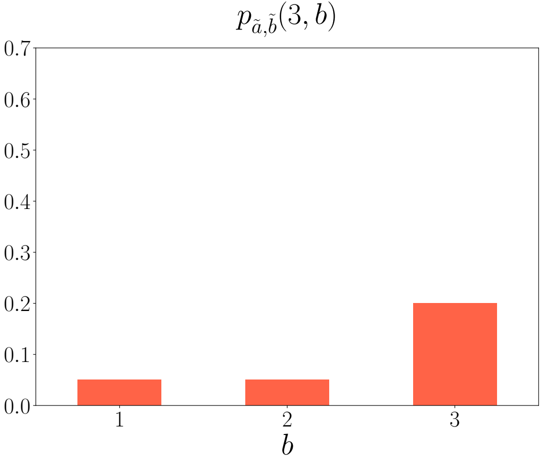
What if we know that $\tilde{a} = 3$?



What if we know that $\tilde{a} = 3$?



Is this a valid pmf?

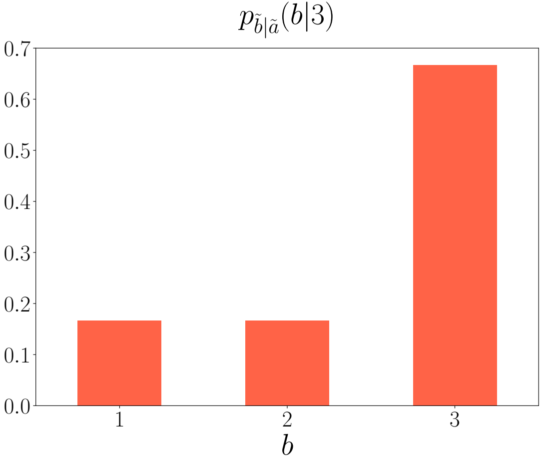


Conditional pmf

The conditional pmf of $\tilde{b}$ given $\tilde{a}$ is

$$\begin{aligned} p_{\tilde{b}\,|\,\tilde{a}}(b\,|\,a) &:= \mathrm{P}\left(\tilde{b} = b\,|\,\tilde{a} = a\right) \\ &= \frac{\mathrm{P}\left(\tilde{a} = a, \tilde{b} = b\right)}{\mathrm{P}\left(\tilde{a} = a\right)} \\ &= \frac{p_{\tilde{a},\tilde{b}}(a, b)}{p_{\tilde{a}}(a)} \end{aligned}$$

Conditional pmf given $\tilde{a} = 3$



Conditional pmf is a valid pmf

Nonnegative? Yes

$$\sum_{b \in B} p_{\tilde{b} \,|\, \tilde{a}}(b \,|\, a) = \frac{\sum_{b \in B} p_{\tilde{a},\tilde{b}}(a, b)}{p_{\tilde{a}}(a)}$$
$$= \frac{p_{\tilde{a}}(a)}{p_{\tilde{a}}(a)}$$
$$= 1$$

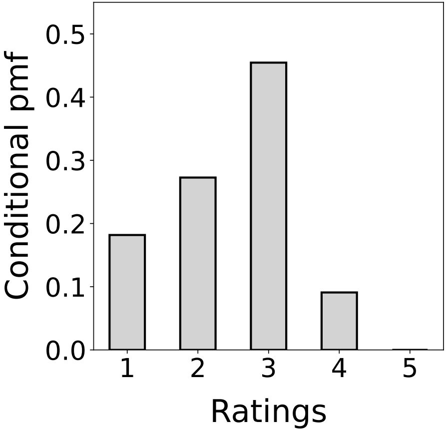
What about $\sum_{a \in A} p_{\tilde{b} \,|\, \tilde{a}}(b \,|\, a)$?

Mission Impossible = 1?

Independence Day (columns) / Mission Impossible (rows)

	1	**2**	**3**	**4**	**5**
1	0.6	1	1.6	0.3	0
2	1	3.8	5.7	3.5	1.6
3	1.6	4.5	11.8	13.1	5.4
4	1.9	4.8	6.4	15	6.1
5	0	0	1.3	3.8	5.4

Conditional pmf



Marginal and conditional pmfs

Marginal and conditional pmfs

Conditional joint pmf of several variables

4-dimensional random vector $\tilde{x}$

$$
\begin{aligned}
& p_{\tilde{x}[2],\tilde{x}[3] \mid \tilde{x}[1],\tilde{x}[4]}(b, c \mid a, d) \\
& \qquad = \mathrm{P}\left(\tilde{x}[2] = b, \tilde{x}[3] = c \mid \tilde{x}[1] = a, \tilde{x}[4] = d\right) \\
& \qquad = \frac{p_{\tilde{x}}(a, b, c, d)}{p_{\tilde{x}[1],\tilde{x}[4]}(a, d)}
\end{aligned}
$$

Chain rule for discrete random variables

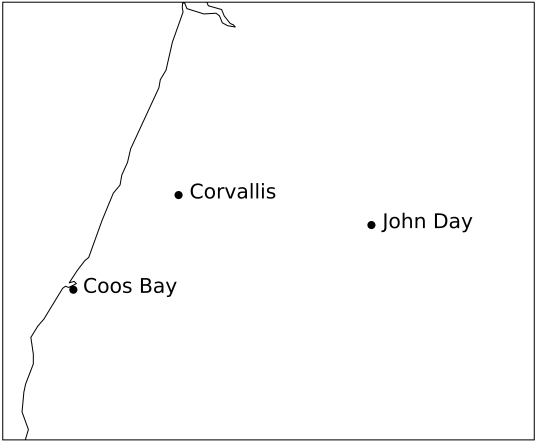
$$
\begin{aligned}
p_{\tilde{a},\tilde{b}}(a,b) &= p_{\tilde{a}}(a)\, p_{\tilde{b}\,|\,\tilde{a}}(b\,|\,a) \\
&= p_{\tilde{b}}(b)\, p_{\tilde{a}\,|\,\tilde{b}}(a\,|\,b)
\end{aligned}
$$

Chain rule for discrete random vectors

$$p_{\tilde{x}}(x) = p_{\tilde{x}[1]}(x[1]) \prod_{i=1}^{n} p_{\tilde{x}[i] \,|\, \tilde{x}[1],\ldots,\tilde{x}[i-1]}(x[i] \,|\, x[1],\ldots,x[i-1])$$

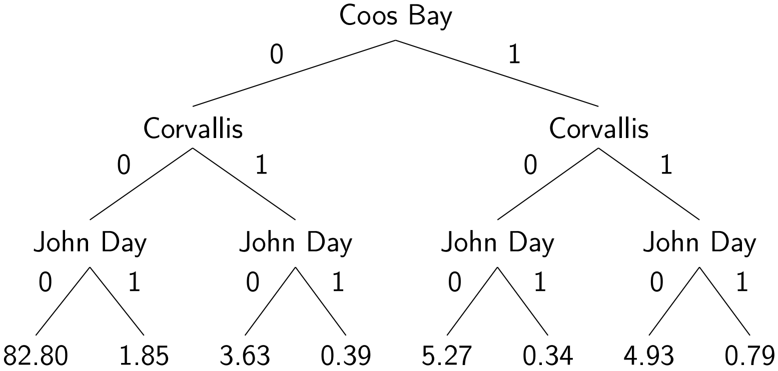
Any order works!

Precipitation in Oregon

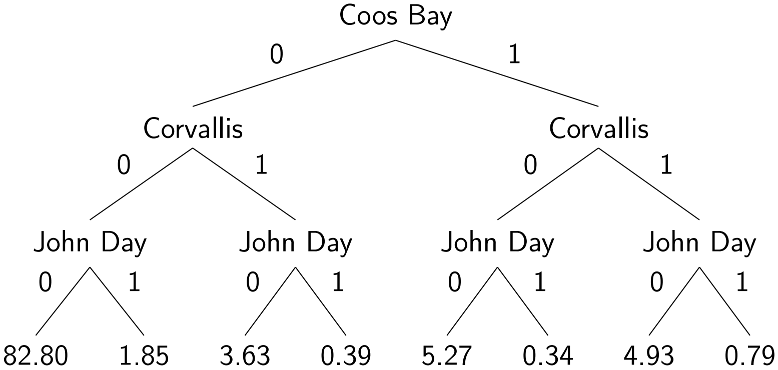


Goal: Model precipitation in Coos Bay, Corvallis, John Day

Precipitation in Oregon (%)



John Day = 0?



Coos Bay
0
1
Corvallis
Corvallis
0
1
0
1
John Day
John Day
John Day
John Day
0
1
0
1
0
1
0
1
82.80
1.85
3.63
0.39
5.27
0.34
4.93
0.79

Conditional joint pmf

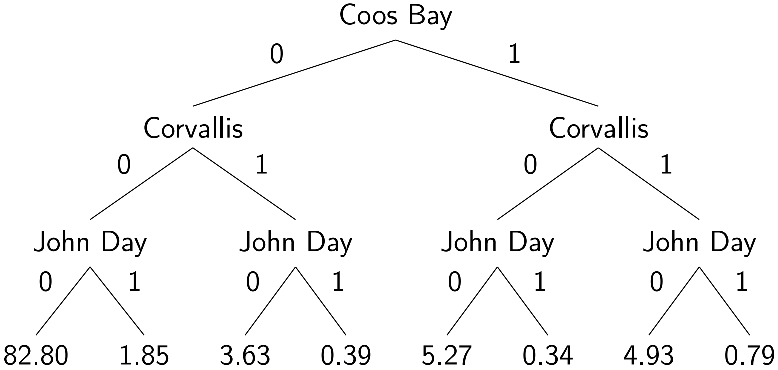
Coos Bay \ Corvallis	**0**	**1**
0	84.70	4.02
1	5.62	5.72

Marginal

Coos Bay \ Corvallis	**0**	**1**
0	85.68	3.76
1	5.46	5.10

John Day = 0

John Day = 1?


Coos Bay
0
1
Corvallis
Corvallis
0
1
0
1
John Day
John Day
John Day
John Day
0
1
0
1
0
1
0
1
82.80
1.85
3.63
0.39
5.27
0.34
4.93
0.79

Conditional joint pmf

Coos Bay \ Corvallis	**0**	**1**
0	84.70	4.02
1	5.62	5.72

Marginal

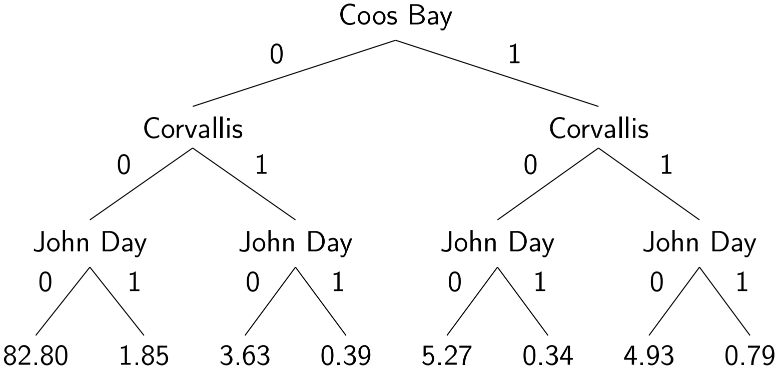
Coos Bay \ Corvallis	**0**	**1**
0	85.68	3.76
1	5.46	5.10

John Day = 0

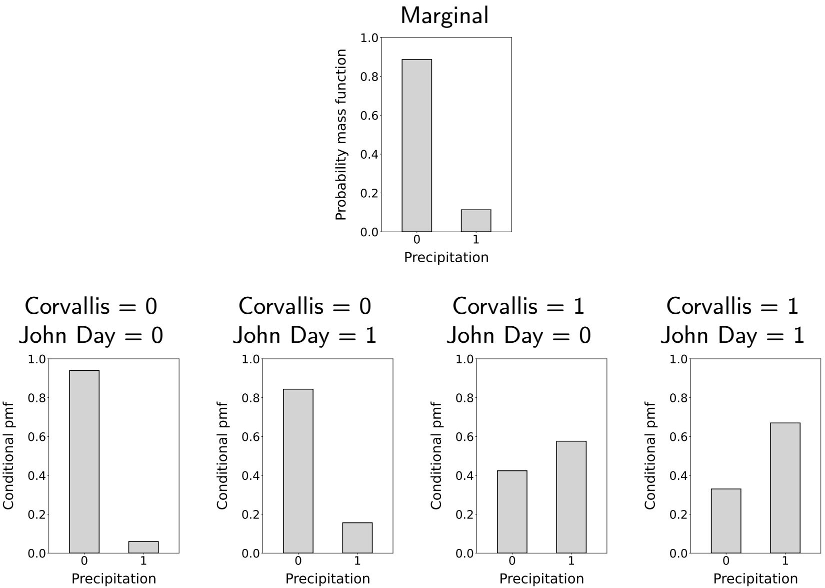
Coos Bay \ Corvallis	**0**	**1**
0	54.92	11.53
1	10.17	23.39

John Day = 1

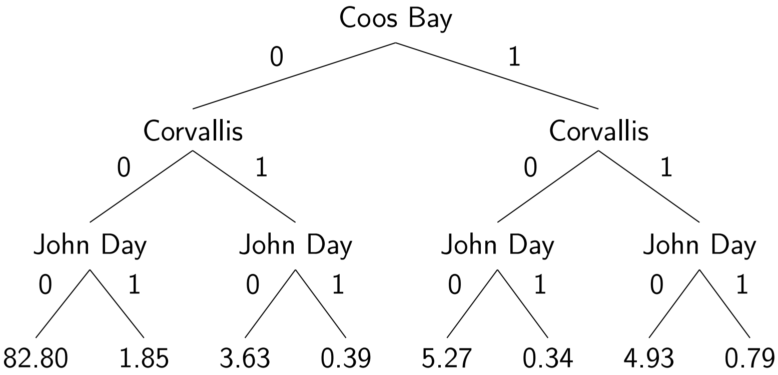
Corvallis = 0, John Day = 0?



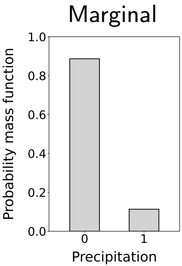
Conditional pmfs

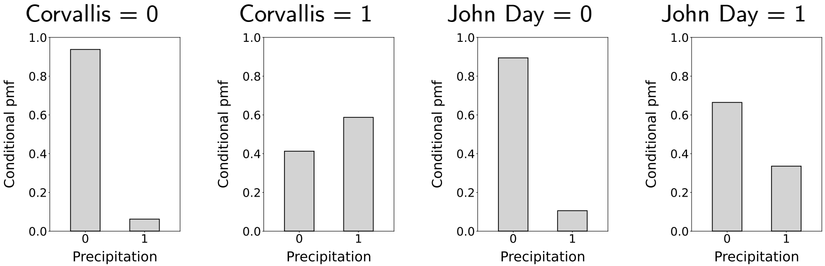


Coos Bay given only Corvallis = 0?

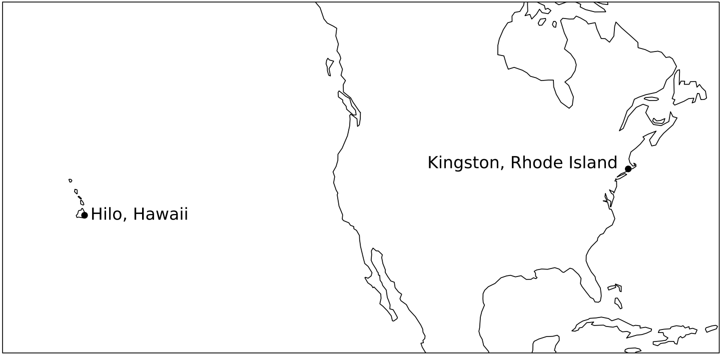


Conditional pmfs



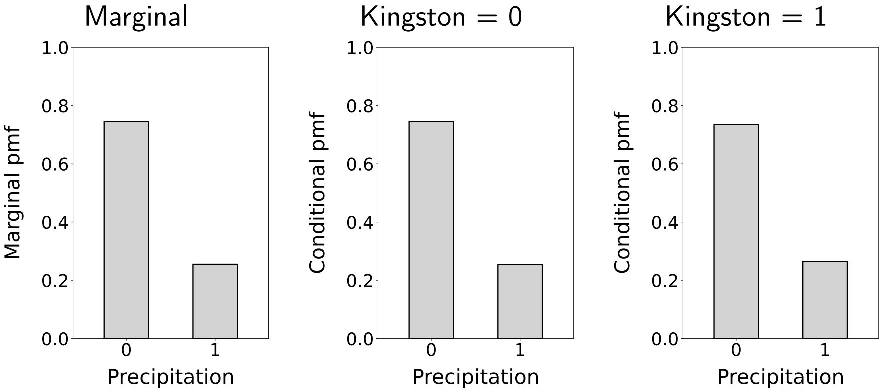


Precipitation in Kingston and Hilo

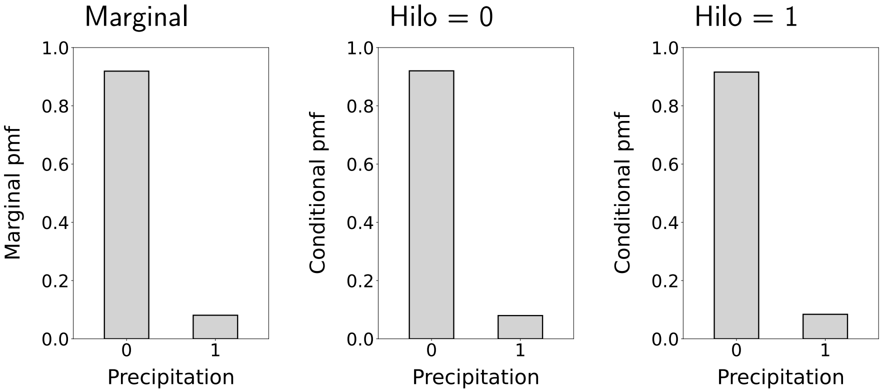


Goal: Model precipitation in Kingston and Hilo

Marginal and conditional pmfs



Marginal and conditional pmfs



Intuition

Two random variables $\tilde{a}$ and $\tilde{b}$ are independent if our uncertainty about $\tilde{a}$ does not change when information about $\tilde{b}$ is revealed

Independence of two events

Two events A, B are independent if

$$\mathrm{P}(A \mid B) = \mathrm{P}(A)$$

or equivalently

$$\mathrm{P}(A \cap B) = \mathrm{P}(A)\,\mathrm{P}(B)$$

Independence

$\tilde{a}$ and $\tilde{b}$ are independent if for any a and b

$$p_{\tilde{a} \mid \tilde{b}}(a \mid b) = \mathrm{P}(\tilde{a} = a \mid \tilde{b} = b) = \mathrm{P}(\tilde{a} = a) = p_{\tilde{a}}(a)$$

Equivalently,

$$p_{\tilde{a},\tilde{b}}(a, b) = p_{\tilde{a}}(a) p_{\tilde{b}}(b)$$

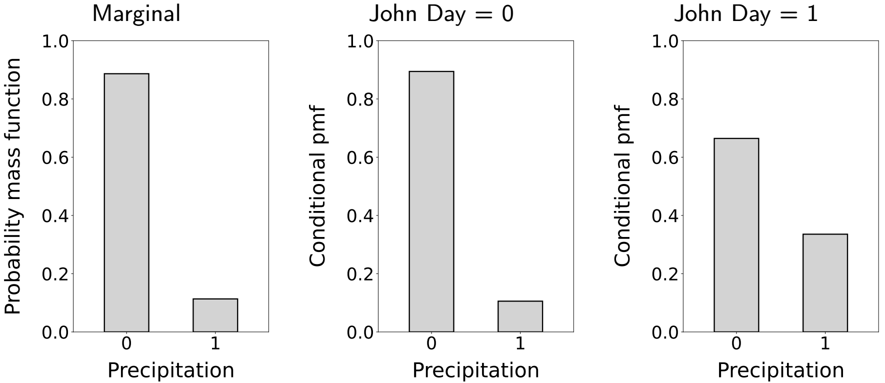
Independence

The d entries $\tilde{x}[1]$, $\tilde{x}[2]$, ..., $\tilde{x}[d]$ in a discrete random vector $\tilde{x}$ are mutually independent if and only if

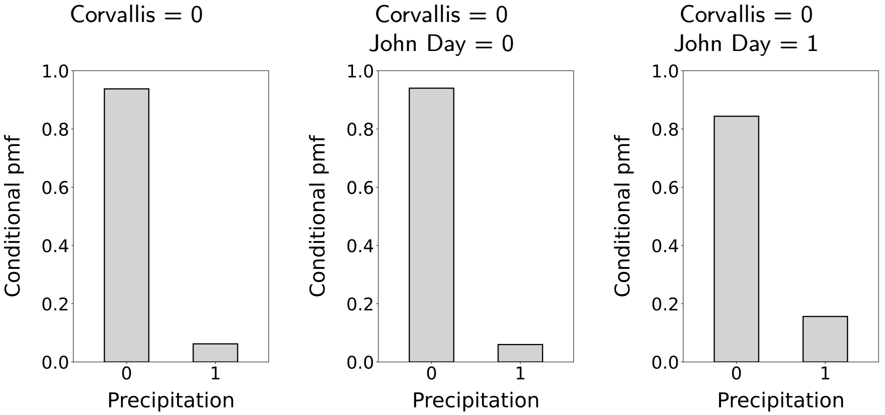
$$p_{\tilde{x}}(x) = \prod_{i=1}^{d} p_{\tilde{x}[i]}(x[i])$$

for all possible values of the entries

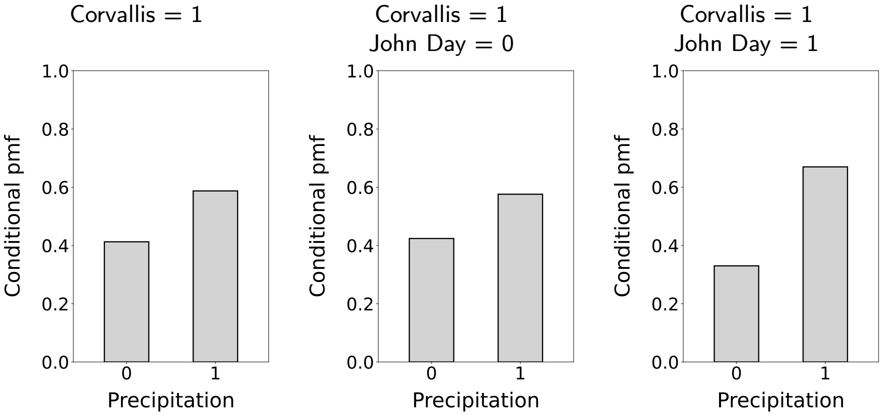
Coos Bay and John Day



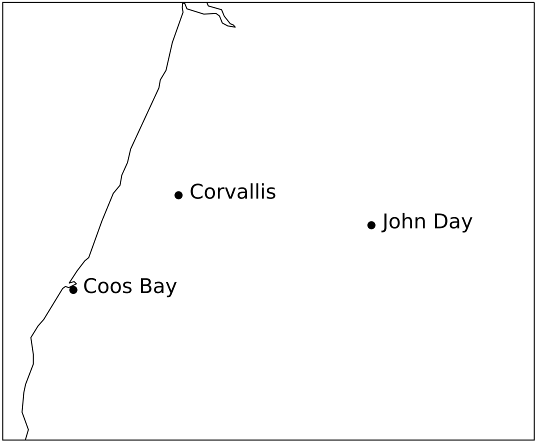
Coos Bay given Corvallis and John Day



Coos Bay given Corvallis and John Day



Precipitation in Oregon



Conditional independence

Events A, B are conditionally independent given C if

$$\mathrm{P}(A \mid B, C) = \mathrm{P}(A \mid C)$$

or equivalently

$$\mathrm{P}(A \cap B \mid C) = \mathrm{P}(A \mid C)\,\mathrm{P}(B \mid C)$$

Conditional independence does not imply independence or vice versa!

Conditional independence

Two random variables $\tilde{a}$ and $\tilde{b}$ are conditionally independent given $\tilde{c}$ if our uncertainty about $\tilde{a}$ does not change when $\tilde{b}$ is revealed, as long as the value of $\tilde{c}$ is known

Conditional independence

$\tilde{a}$ and $\tilde{b}$ are conditionally independent given $\tilde{c}$ if

$$p_{\tilde{a},\tilde{b}\,|\,\tilde{c}}(a, b\,|\,c) = p_{\tilde{a}\,|\,\tilde{c}}(a\,|\,c)\, p_{\tilde{b}\,|\,\tilde{c}}(b\,|\,c) \qquad \text{for all } a, b, c$$

What have we learned?

How to compute conditional pmfs

Definition of independence

Definition of conditional independence