Conditional Distributions of Continuous Random Variables

Probability and Statistics for Data Science

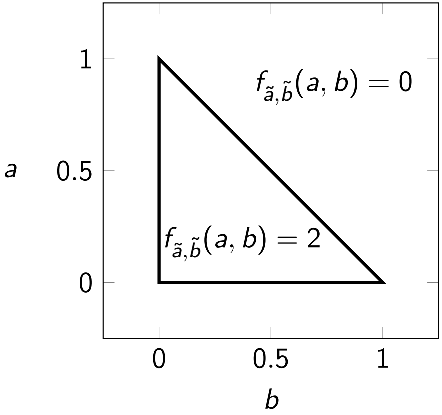
Carlos Fernandez-Granda

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

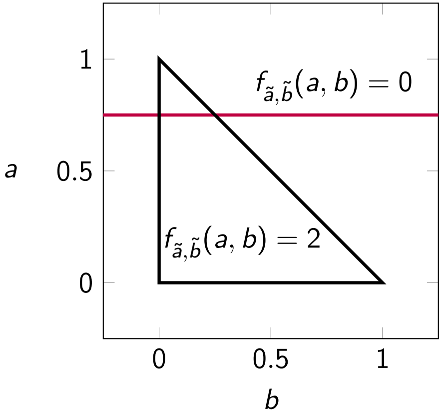
Motivation

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

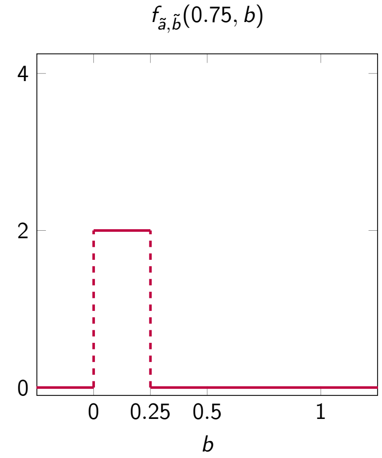
What if we know that $\tilde{a} = 0.75$?



What if we know that $\tilde{a} = 0.75$?



Is this a valid pdf?



Conditional probability density

$$\lim_{\epsilon_1 \to 0} \frac{\mathrm{P}\left(b - \epsilon_1 < \tilde{b} \leq b \,|\, \tilde{a} = a\right)}{\epsilon_1}$$

Probability of event $\tilde{a} = a$? Zero!

Conditional probabilities given $\tilde{a} = a$ are not well defined...

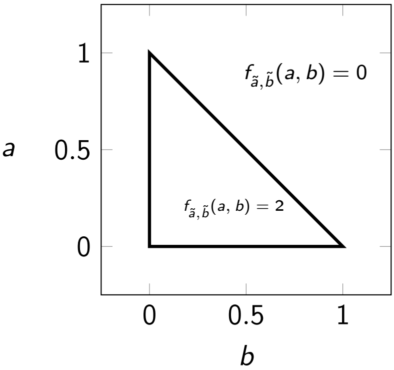
Conditional probability density

$$f_{\tilde{b}\,|\,a-\epsilon_2<\tilde{a}\leq a}(b) := \lim_{\epsilon\to 0} \frac{\mathrm{P}\left(b-\epsilon_1 < \tilde{b} \leq b \,|\, a-\epsilon_2 < \tilde{a} \leq a\right)}{\epsilon_1}$$

Setting $\epsilon = \epsilon_1 = \epsilon_2$

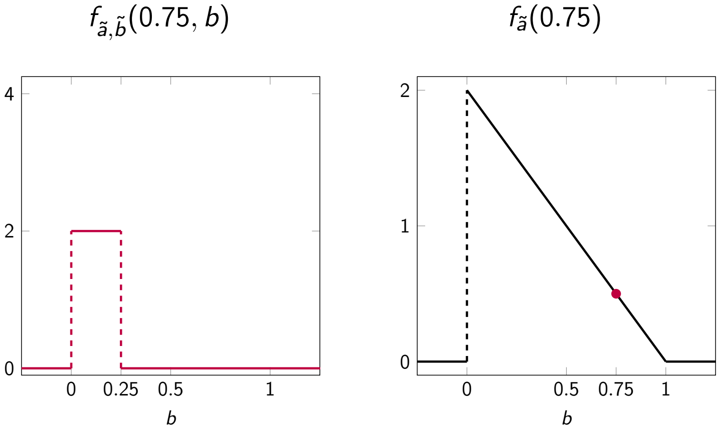
$$
\begin{aligned}
f_{\tilde{b}\,|\,\tilde{a}}(b\,|\,a) &= \lim_{\epsilon\to 0} f_{\tilde{b}\,|\,a-\epsilon<\tilde{a}\leq a}(b) \\
&= \lim_{\epsilon\to 0} \frac{\mathrm{P}\left(b-\epsilon<\tilde{b}\leq b\,|\,a-\epsilon<\tilde{a}\leq a\right)}{\epsilon} \\
&= \lim_{\epsilon\to 0} \frac{1}{\epsilon}\frac{\mathrm{P}\left(b-\epsilon<\tilde{b}\leq b, a-\epsilon<\tilde{a}\leq a\right)}{\mathrm{P}\left(a-\epsilon<\tilde{a}\leq a\right)} \\
&= \frac{\lim_{\epsilon\to 0}\frac{\mathrm{P}\left(b-\epsilon<\tilde{b}\leq b, a-\epsilon<\tilde{a}\leq a\right)}{\epsilon^2}}{\lim_{\epsilon\to 0}\frac{\mathrm{P}\left(a-\epsilon<\tilde{a}\leq a\right)}{\epsilon}} \\
&= \frac{f_{\tilde{a},\tilde{b}}\left(a,b\right)}{f_{\tilde{a}}\left(a\right)}
\end{aligned}
$$

Marginal pdf



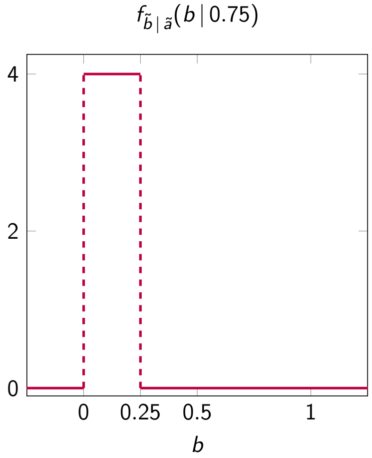
$$f_{\tilde{a}}(a) = \int_{b=-\infty}^{\infty} f_{\tilde{a},\tilde{b}}(a,b)\,\mathrm{d}b$$
$$= \int_{b=0}^{1-a} 2\,\mathrm{d}b = 2(1-a)$$

Conditional pdf



$$f_{\tilde{b}\,|\,\tilde{a}}(b\,|\,a) = \frac{f_{\tilde{a},\tilde{b}}(a, b)}{f_{\tilde{a}}(a)} = \frac{1}{1-a} \qquad b \in [0, 1-a]$$

Conditional pdf



Conditional pdf

Conditional pdf of $\tilde{b}$ given $\tilde{a}$

$$f_{\tilde{b}\,|\,\tilde{a}}\left(b\,|\,a\right) := \frac{f_{\tilde{a},\tilde{b}}\left(a,b\right)}{f_{\tilde{a}}\left(a\right)} \quad \text{if } f_{\tilde{a}}\left(x\right) > 0$$

Conditional pdf of $\tilde{x}[i]$ given $\tilde{x}[j] = a_j$ for $j \neq i$

$$\begin{aligned} & f_{\tilde{x}[i]\,|\,\tilde{x}[1],\ldots,\tilde{x}[i-1],\tilde{x}[i+1],\ldots,\tilde{x}[d]}\left(b\,|\,a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_d\right) \\ & = \frac{f_{\tilde{x}}\left(a_1,\ldots,a_{i-1},b,a_{i+1},\ldots,a_d\right)}{f_{\tilde{x}[1],\ldots,\tilde{x}[i-1],\tilde{x}[i+1],\ldots,\tilde{x}[d]}\left(a_1,\ldots,a_{i-1},a_{i+1},\ldots,a_d\right)} \end{aligned}$$

Conditional pdf

Conditional joint pdf of $\tilde{x}[2]$ and $\tilde{x}[3]$ given $\tilde{x}[1]$ and $\tilde{x}[4]$

$$f_{\tilde{x}[2],\tilde{x}[3]\,|\,\tilde{x}[1],\tilde{x}[4]}(b, c \mid a, d) = \frac{f_{\tilde{x}}(a, b, c, d)}{f_{\tilde{x}[1],\tilde{x}[4]}(a, d)}$$

The conditional pdf is a valid pdf

$$
\begin{aligned}
\int_{b=-\infty}^{\infty} f_{\tilde{b} \,|\, \tilde{a}}(b \,|\, a) \, \mathrm{d}b &= \frac{\int_{b=-\infty}^{\infty} f_{\tilde{a},\tilde{b}}(a, b) \, \mathrm{d}b}{f_{\tilde{a}}(a)} \\
&= \frac{f_{\tilde{a}}(a)}{f_{\tilde{a}}(a)} \\
&= 1
\end{aligned}
$$

Chain rule

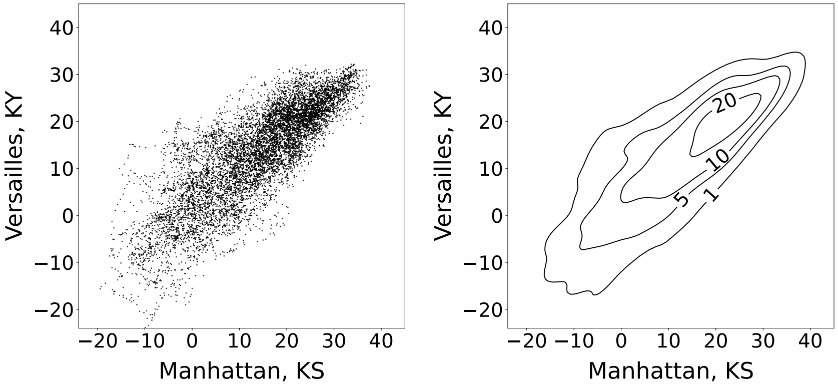
$$\begin{aligned} f_{\tilde{a},\tilde{b}}(a,b) &= f_{\tilde{a}}(a)\, f_{\tilde{b}\,|\,\tilde{a}}(b\,|\,a) \\ &= f_{\tilde{b}}(b)\, f_{\tilde{a}\,|\,\tilde{b}}(a\,|\,b) \end{aligned}$$

Chain rule

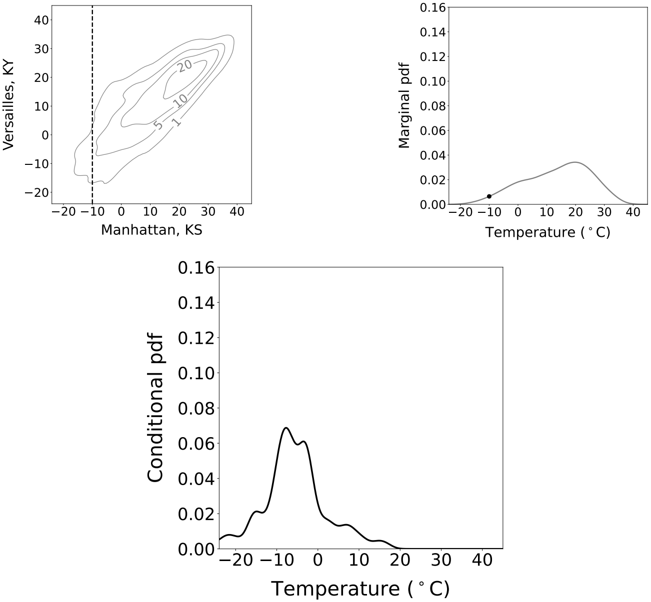
$$f_{\tilde{x}}(x) = f_{\tilde{x}[1]}(x[1]) \prod_{i=1}^{n} f_{\tilde{x}[i] \mid \tilde{x}[1],\ldots,\tilde{x}[i-1]}(x[i] \mid x[1], \ldots, x[i-1])$$

Any order works!

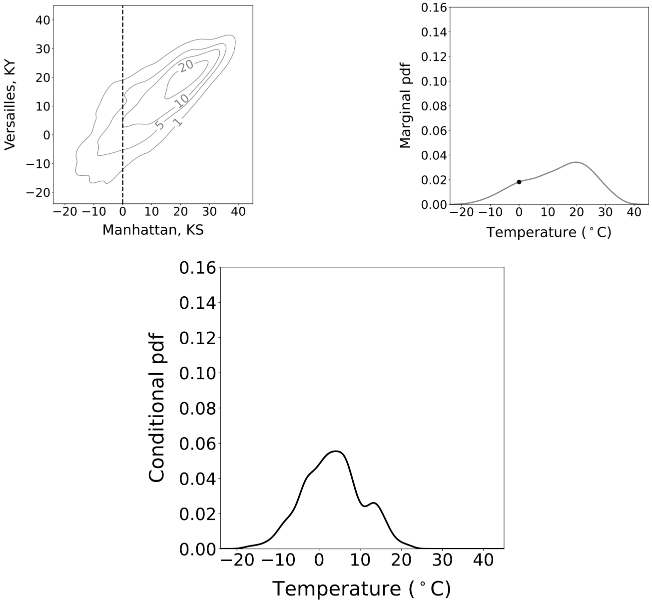
Temperature



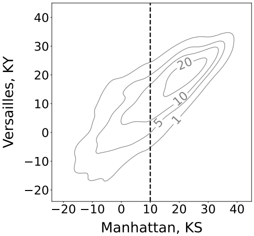
-10° in Manhattan

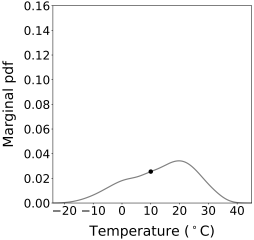


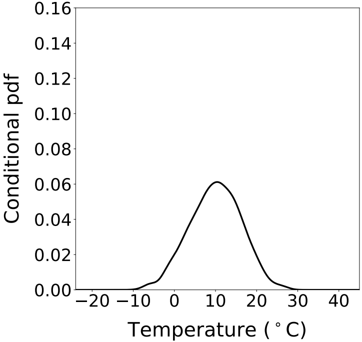
0° in Manhattan



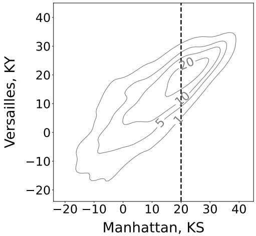
10° in Manhattan

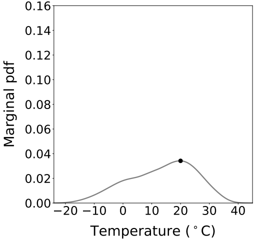


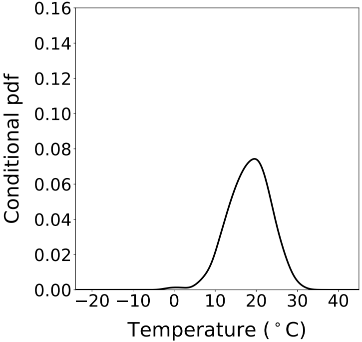




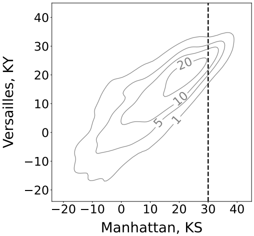
20° in Manhattan

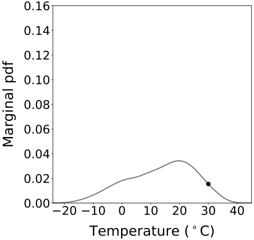


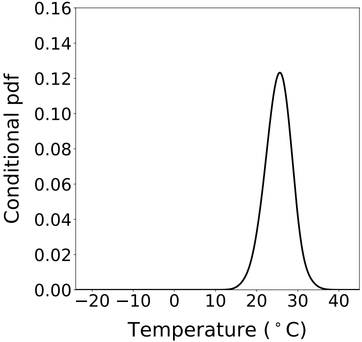




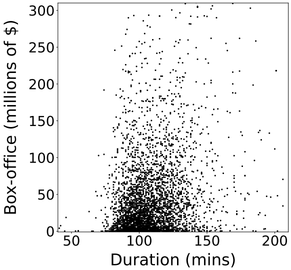
30° in Manhattan

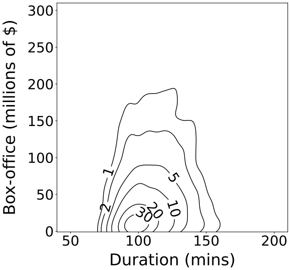


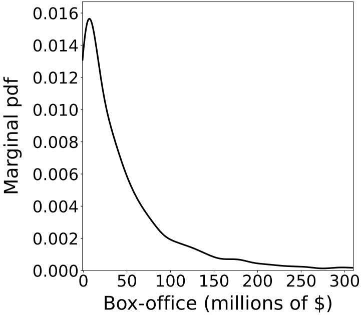




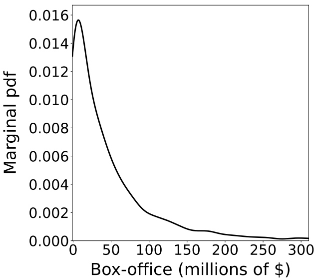
Movie length and box office

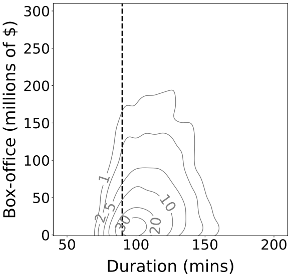


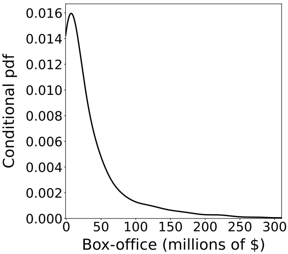




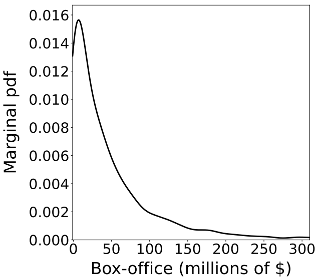
Duration = 90 mins

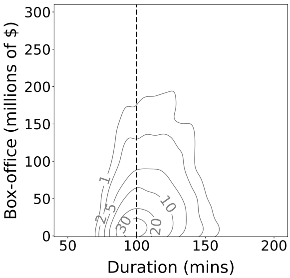


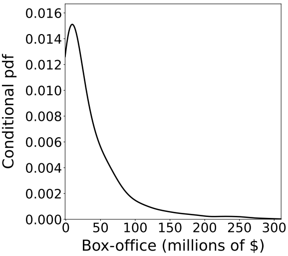




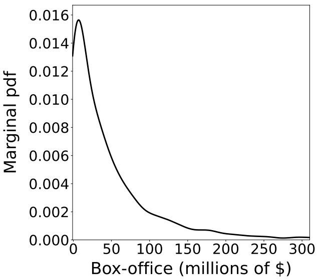
Duration = 100 mins

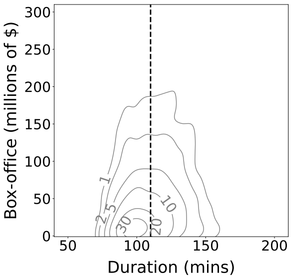


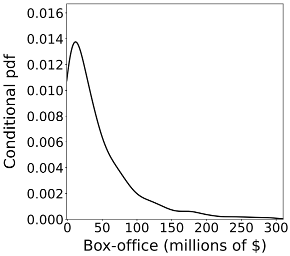




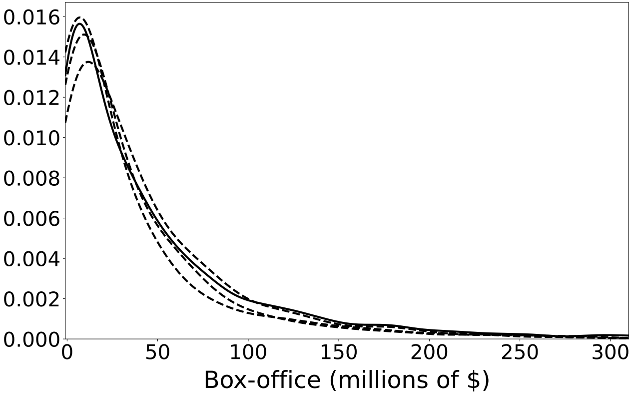
Duration = 110 mins







Marginal and conditional pdfs



Independence

The random variables $\tilde{a}$ and $\tilde{b}$ are independent if for any Borel set S and any b

$$P(\tilde{a} \in S \,|\, \tilde{b} = b) = \mathrm{P}(\tilde{a} \in S)$$

Equivalently,

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

$$f_{\tilde{a} \,|\, \tilde{b}}(a \,|\, b) = f_{\tilde{a}}(a)$$

Independence

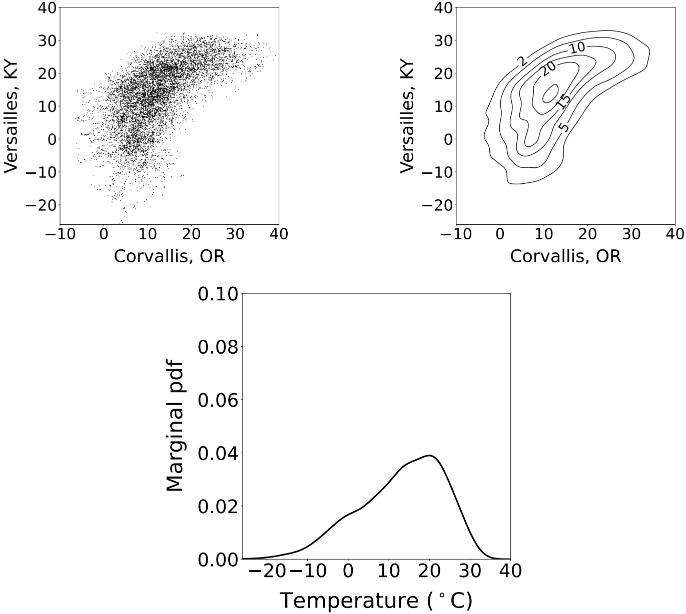
$\tilde{a}$ and $\tilde{b}$ are independent if for any a and b

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

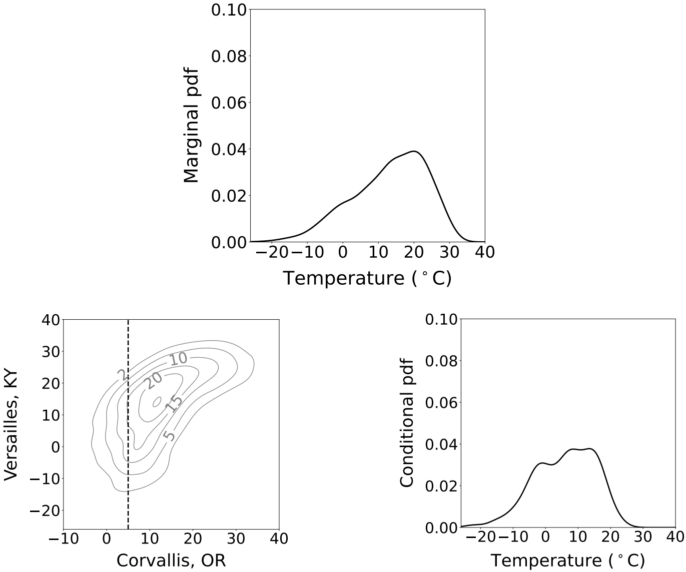
The entries of $\tilde{x}$ are independent if for all x

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

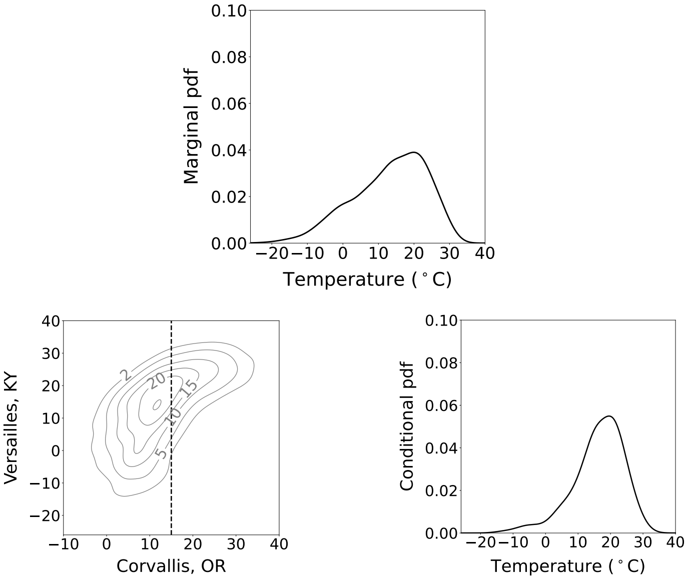
Temperature



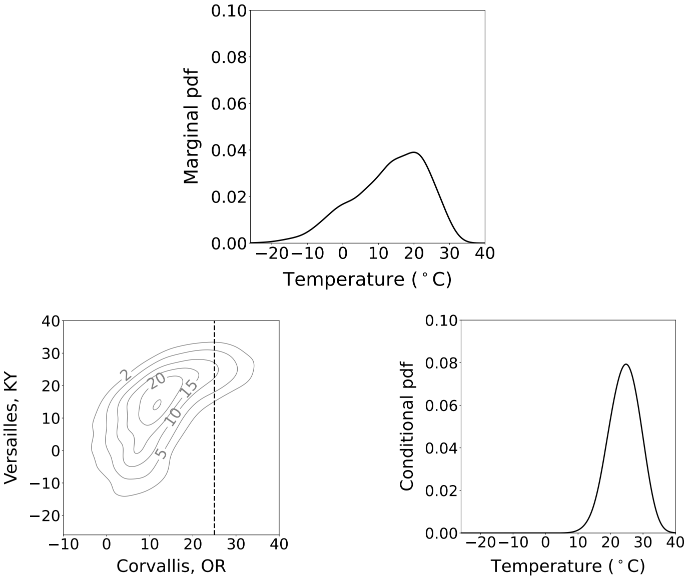
Corvallis = 5°C



Corvallis = 15°C



Corvallis = 25°C

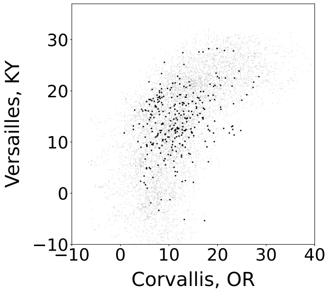


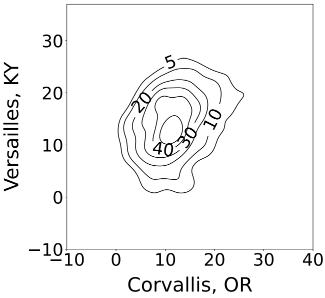
Let us condition on Manhattan

Versailles ($\tilde{v}$) and Corvallis ($\tilde{c}$) given Manhattan ($\tilde{m}$)

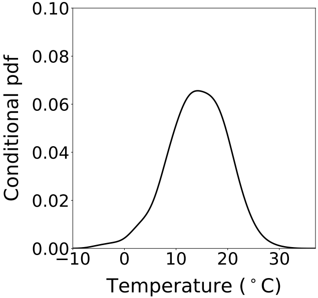
$$f_{\tilde{v},\tilde{c}\,|\,\tilde{m}}(v, c \,|\, t) = \frac{f_{\tilde{v},\tilde{c},\tilde{m}}(v, c, t)}{f_{\tilde{m}}(t)}$$

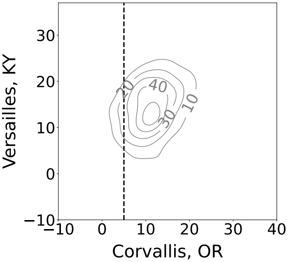
Manhattan = 15°C

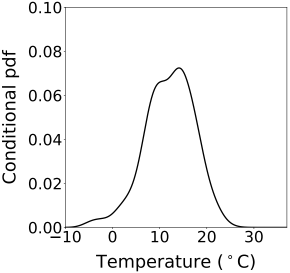




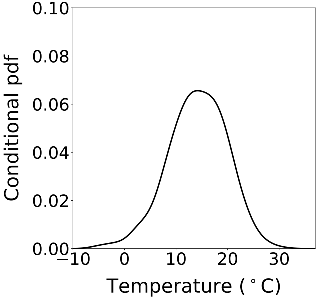
Manhattan = 15°C, Corvallis = 5°C

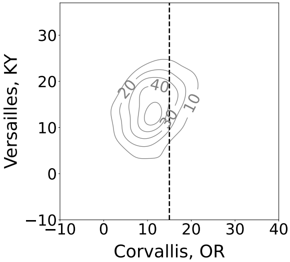


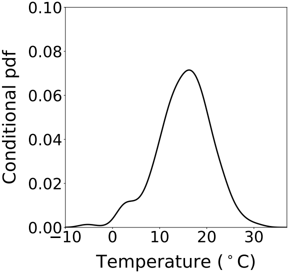




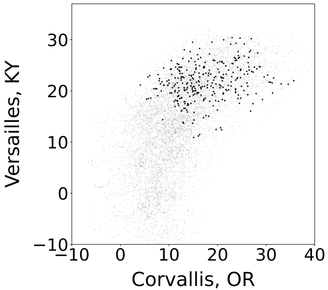
Manhattan = 15°C, Corvallis = 15°C

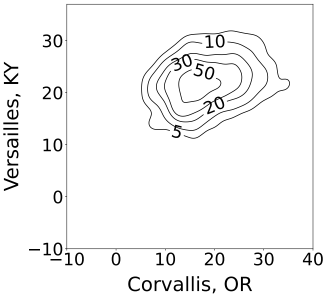




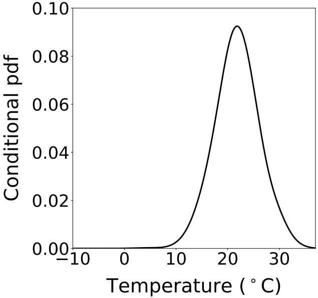


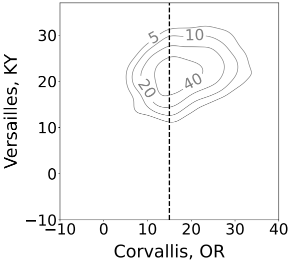
Manhattan = 25°C

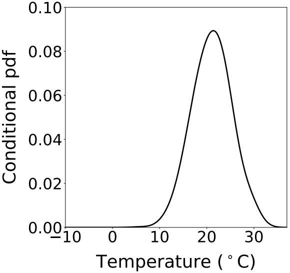




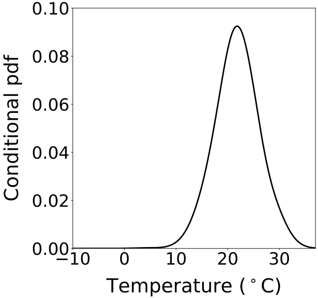
Manhattan = 25°C, Corvallis = 15°C

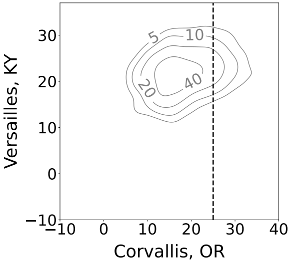


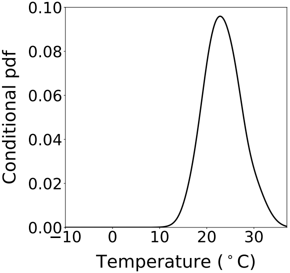




Manhattan = 25°C, Corvallis = 25°C







Corvallis, Manhattan, Versailles

Conditional independence

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

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

$\tilde{x}[1]$, $\tilde{x}[2]$, ..., $\tilde{x}[d_1]$ are conditionally independent given $\tilde{y}$ if and only if

$$f_{\tilde{x}\,|\,\tilde{y}}(x\,|\,y) = \prod_{i=1}^{d} f_{\tilde{x}[i]\,|\,\tilde{y}}(x[i]\,|\,y), \quad \text{for all } x \in \mathbb{R}^{d_1}, y \in \mathbb{R}^{d_2}$$

What have we learned?

How to compute conditional pdfs

Definition of independence / conditional independence for continuous random variables