Marginal Distributions of Continuous Random Variables

Probability and Statistics for Data Science

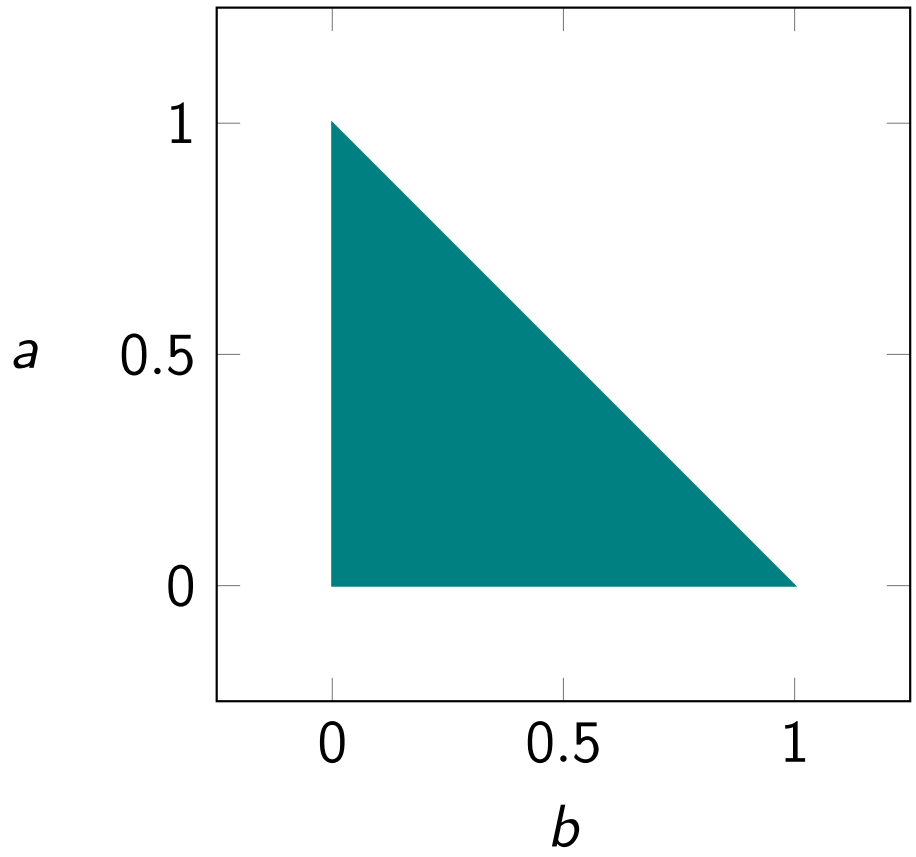
Carlos Fernandez-Granda

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

Goal

In a model with many variables, how do we characterize behavior of individual variables?

Triangle lake: Joint pdf



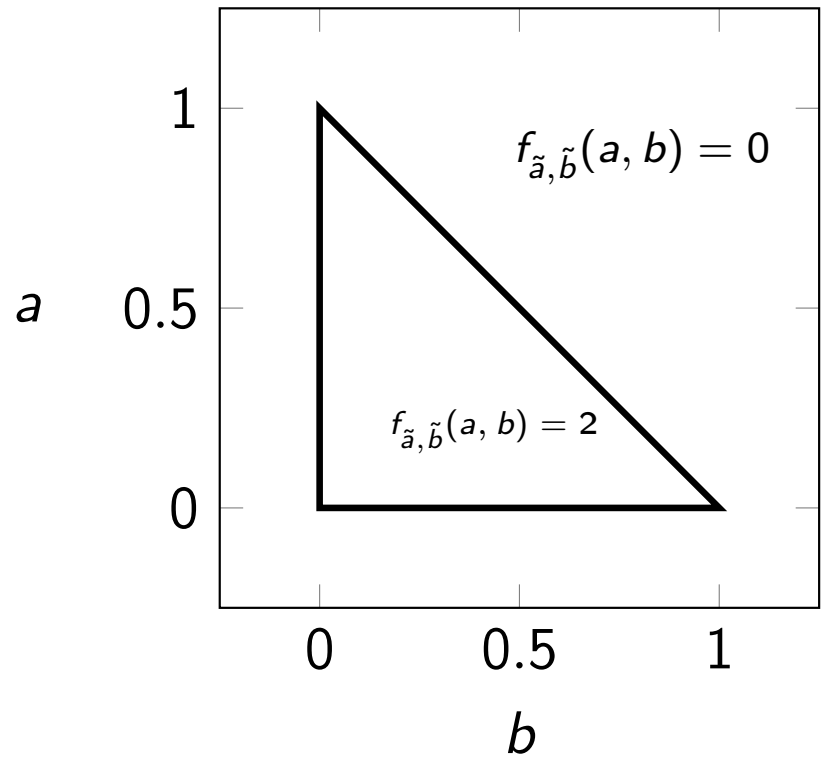
What if we only care about $\tilde{a}$?

Marginal pdf

$$
\begin{aligned}
F_{\tilde{a}}(a) &= \mathrm{P}(\tilde{a} \leq a) \\
&= \int_{u=-\infty}^{a} \int_{b=-\infty}^{\infty} f_{\tilde{a},\tilde{b}}(u, b)\, \mathrm{d}b\, \mathrm{d}u
\end{aligned}
$$

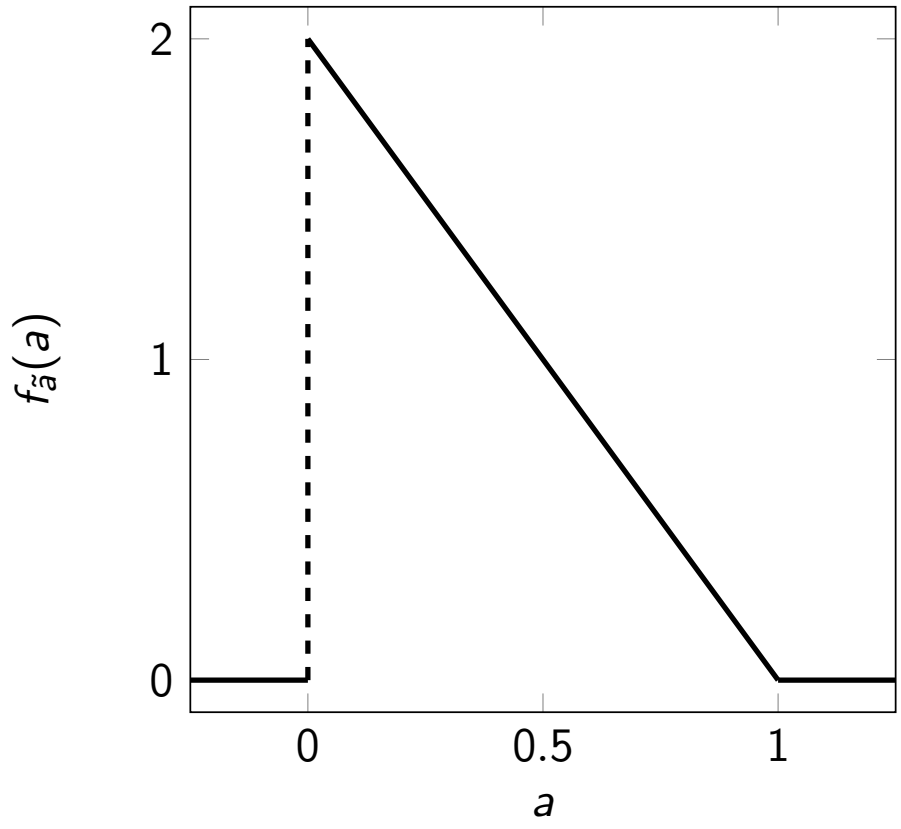
$$
f_{\tilde{a}}(a) = \int_{b=-\infty}^{\infty} f_{\tilde{a},\tilde{b}}(a, b)\ \mathrm{d}b
$$

Marginal pdf



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

Marginal pdf



Marginal pdf

Marginal pdf of $\tilde{a}$

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

Marginal pdf of $\tilde{x}[i]$

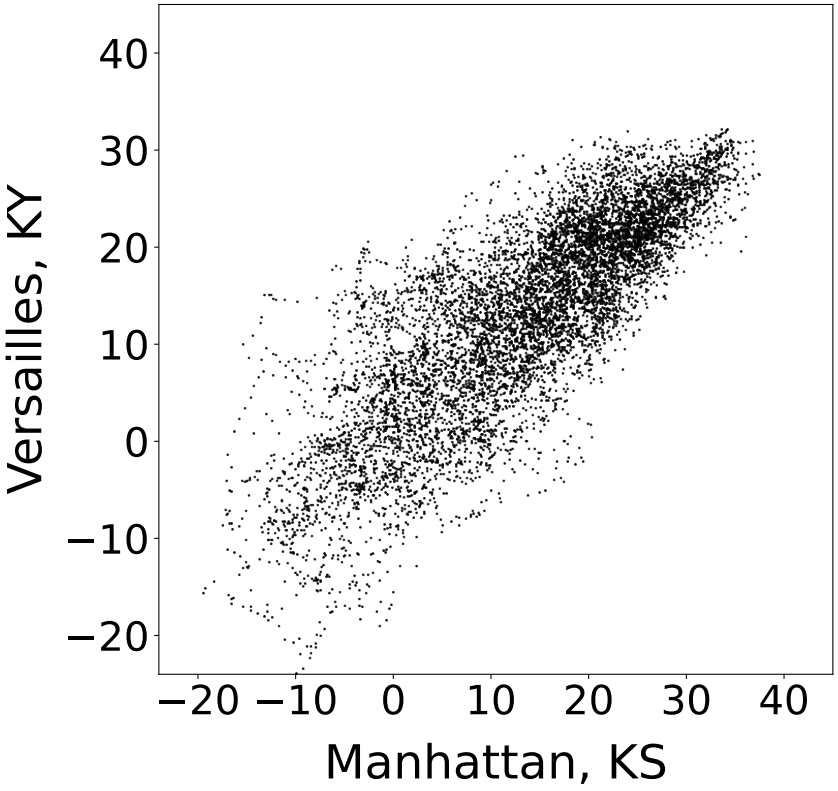
$$f_{\tilde{x}[i]}(a) =$$
$$\int_{x[\mathbf{1}]} \cdots \int_{x[i-\mathbf{1}]} \int_{x[i+\mathbf{1}]} \cdots \int_{x[d]} f_{\tilde{x}}\left(x[\mathbf{1}], \ldots, x[i-\mathbf{1}], a, x[i+\mathbf{1}], \ldots, x[d]\right)\ \mathbf{d}x[\mathbf{1}] \ldots \mathbf{d}x[i-\mathbf{1}]\, \mathbf{d}x[i+\mathbf{1}] \ldots \mathbf{d}x[d]$$

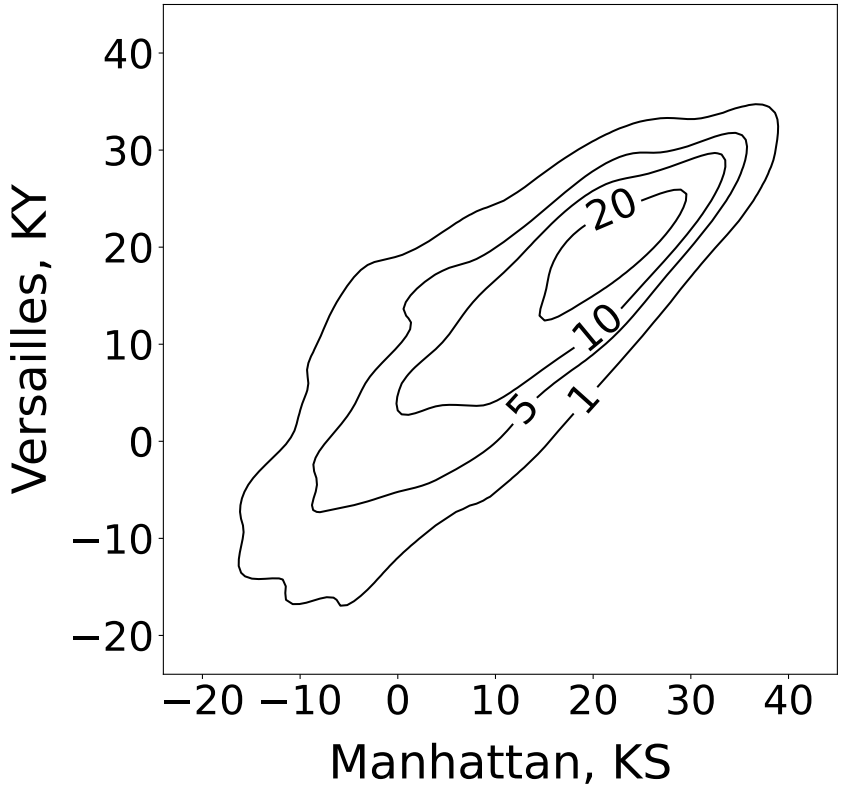
Marginal pdf

Marginal joint pdf of $\tilde{x}[1]$ and $\tilde{x}[4]$

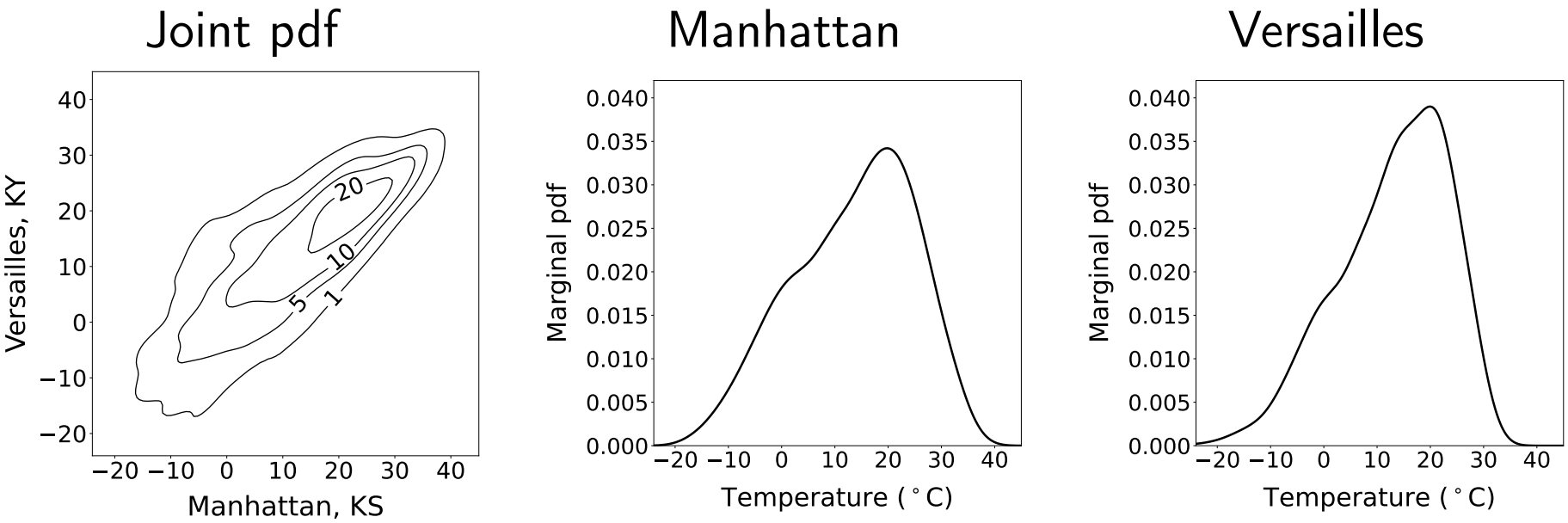
$$f_{\tilde{x}[1],\tilde{x}[4]}(a,d) = \int_{b=-\infty} \int_{c=-\infty} f_{\tilde{x}}(a,b,c,d)\, \mathrm{d}b\, \mathrm{d}c$$

Temperature





Marginal distributions



What have we learned?

Definition of marginal distribution

How to compute it from the joint pdf