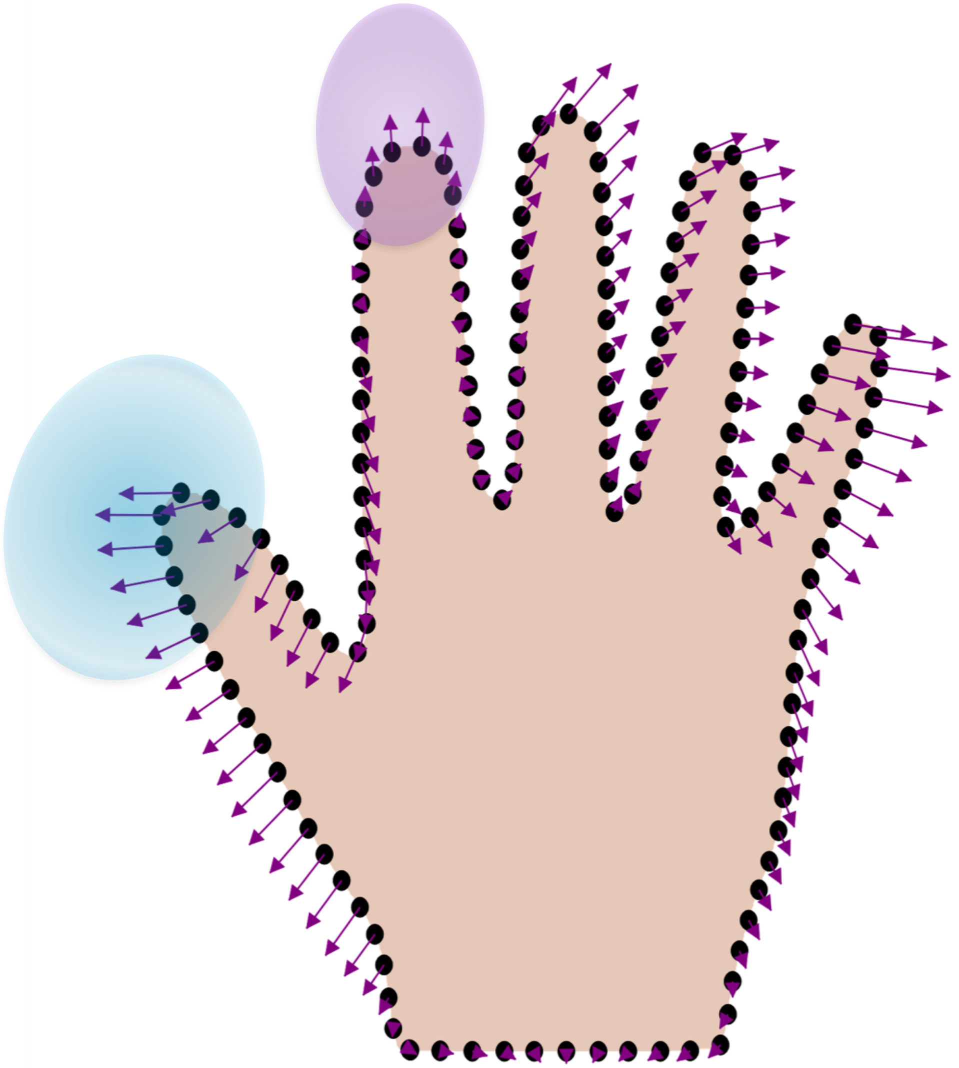
University
of Basel

The marginalization property

Constructing Gaussian processes

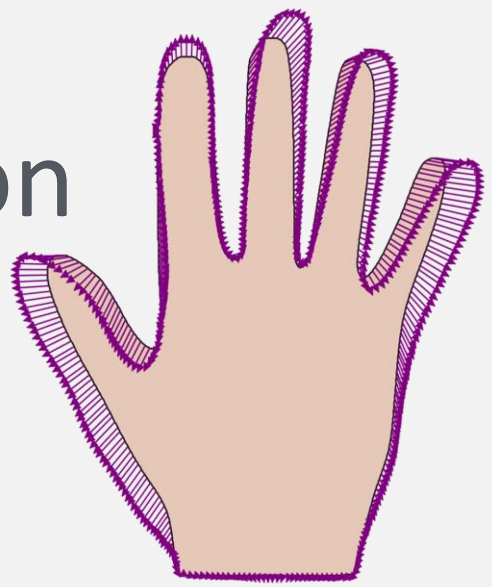
University of Basel



$$N\left(\begin{pmatrix}\mu_1(x_1)\\ \mu_2(x_1)\\ \vdots \\ \mu_1(x_n)\\ \mu_2(x_n)\end{pmatrix}, \begin{pmatrix} k_{11}(x_1,x_1) & k_{12}(x_1,x_1) & \cdots & k_{11}(x_1,x_n) & k_{12}(x_1,x_n)\\ k_{21}(x_1,x_1) & k_{22}(x_1,x_1) & & k_{21}(x_1,x_n) & k_{22}(x_1,x_n)\\ & \vdots & \ddots & & \vdots \\ k_{11}(x_n,x_1) & k_{12}(x_n,x_1) & \cdots & k_{11}(x_n,x_n) & k_{12}(x_n,x_n)\\ k_{21}(x_n,x_1) & k_{22}(x_n,x_1) & & k_{21}(x_n,x_n) & k_{22}(x_n,x_n)\end{pmatrix}\right)$$

Constructing Gaussian processes

Continuous representation

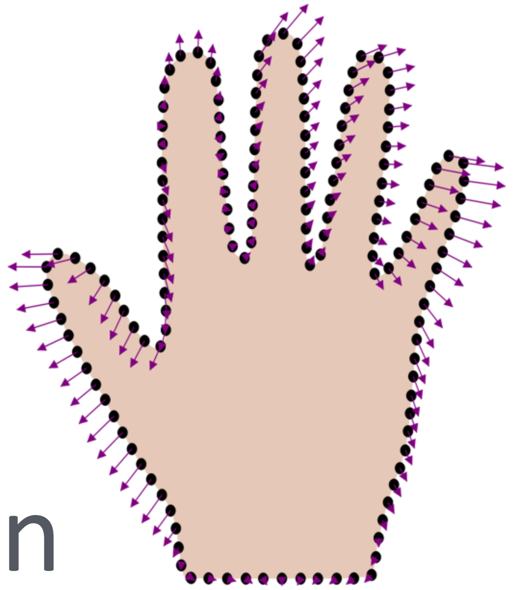


$GP(\mu, k)$

$\mu: \Gamma_R \rightarrow \mathbb{R}^2$

$k: \Gamma_R \times \Gamma_R \rightarrow \mathbb{R}^{2\times 2}$

Discrete representation

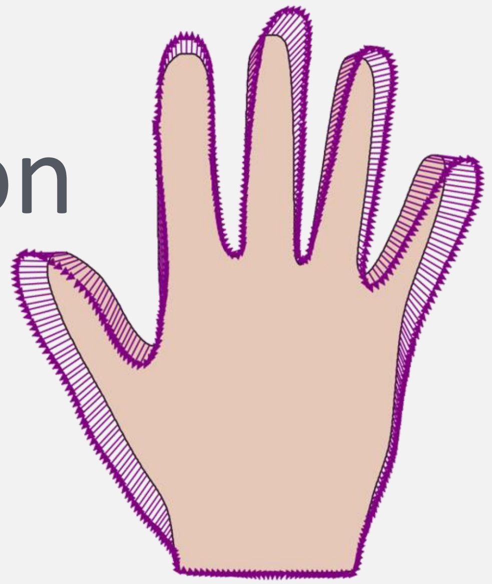


$$N\left(\begin{pmatrix}\mu_1(x_1)\\ \mu_2(x_1)\\ \vdots \\ \mu_1(x_n)\\ \mu_2(x_n)\end{pmatrix}, \begin{pmatrix} k_{11}(x_1,x_1) & k_{12}(x_1,x_1) & \cdots & k_{11}(x_1,x_n) & k_{12}(x_1,x_n) \\ k_{21}(x_1,x_1) & k_{22}(x_1,x_1) & & k_{21}(x_1,x_n) & k_{22}(x_1,x_n) \\ & \vdots & \ddots & & \vdots \\ k_{11}(x_n,x_1) & k_{12}(x_n,x_1) & \cdots & k_{11}(x_n,x_n) & k_{12}(x_n,x_n) \\ k_{21}(x_n,x_1) & k_{22}(x_n,x_1) & & k_{21}(x_n,x_n) & k_{22}(x_n,x_n)\end{pmatrix}\right)$$

Constructing Gaussian processes

University of Basel

Continuous representation



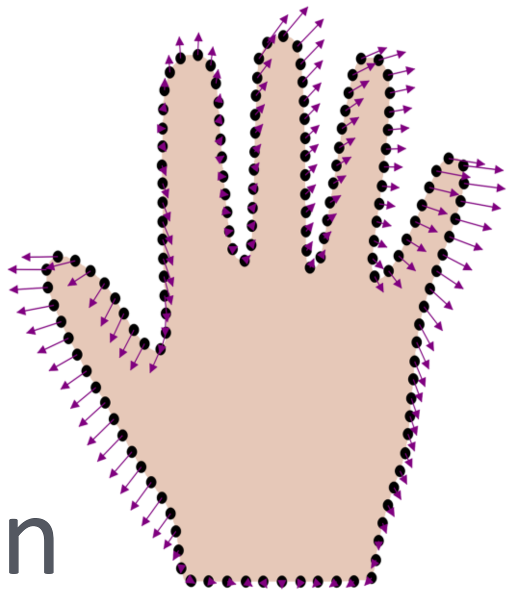
$GP(\mu, k)$

+ Simple and compact definition
+ Can be infinite dimensional
+ Independent of discretization

$\mu: \Gamma_R \to \mathbb{R}^2$

$k: \Gamma_R \times \Gamma_R \to \mathbb{R}^{2\times 2}$

Discrete representation

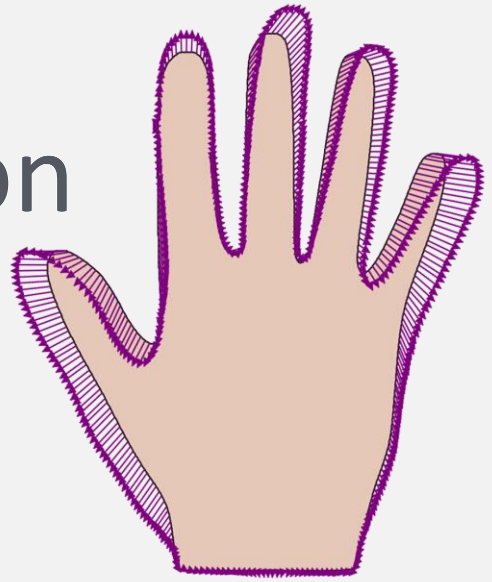


$$N\left(\begin{pmatrix} \mu_1(x_1) \\ \mu_2(x_1) \\ \vdots \\ \mu_1(x_n) \\ \mu_2(x_n) \end{pmatrix}, \begin{pmatrix} k_{11}(x_1,x_1) & k_{12}(x_1,x_1) & \cdots & k_{11}(x_1,x_n) & k_{12}(x_1,x_n) \\ k_{21}(x_1,x_1) & k_{22}(x_1,x_1) & & k_{21}(x_1,x_n) & k_{22}(x_1,x_n) \\ \vdots & & \ddots & & \vdots \\ k_{11}(x_n,x_1) & k_{12}(x_n,x_1) & \cdots & k_{11}(x_n,x_n) & k_{12}(x_n,x_n) \\ k_{21}(x_n,x_1) & k_{22}(x_n,x_1) & & k_{21}(x_n,x_n) & k_{22}(x_n,x_n) \end{pmatrix}\right)$$

Constructing Gaussian processes

University of Basel

Continuous representation



$GP(\mu, k)$

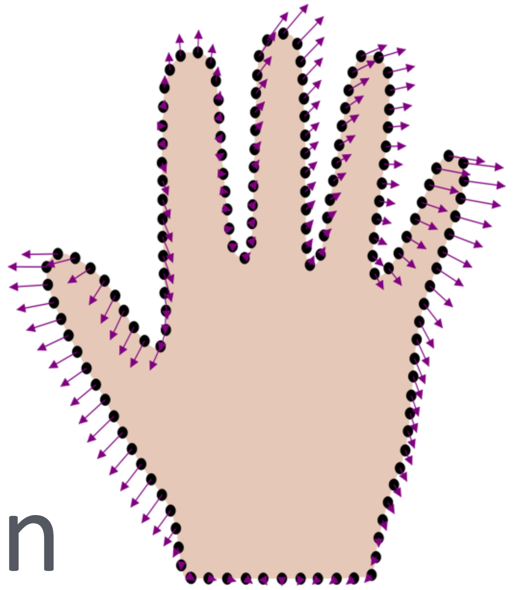
+ Simple and compact definition
+ Can be infinite dimensional
+ Independent of discretization

$\mu: \Gamma_R \to \mathbb{R}^2$

$k: \Gamma_R \times \Gamma_R \to \mathbb{R}^{2\times 2}$

? Computer implementation

Discrete representation



$$N\left(\begin{pmatrix}\mu_1(x_1)\\ \mu_2(x_1)\\ \vdots \\ \mu_1(x_n)\\ \mu_2(x_n)\end{pmatrix}, \begin{pmatrix} k_{11}(x_1,x_1) & k_{12}(x_1,x_1) & \dots & k_{11}(x_1,x_n) & k_{12}(x_1,x_n)\\ k_{21}(x_1,x_1) & k_{22}(x_1,x_1) & & k_{21}(x_1,x_n) & k_{22}(x_1,x_n)\\ & \vdots & \ddots & \vdots & \\ k_{11}(x_n,x_1) & k_{12}(x_n,x_1) & \dots & k_{11}(x_n,x_n) & k_{12}(x_n,x_n)\\ k_{21}(x_n,x_1) & k_{22}(x_n,x_1) & & k_{21}(x_n,x_n) & k_{22}(x_n,x_n)\end{pmatrix}\right)$$

Marginalization property

University of Basel

Let $X = (x_1, \dots, x_n)$ and $Y = (y_1, \dots, y_m)$

be jointly normal distributed random variables

$$p(X, Y) = N\left(\begin{pmatrix}\mu_X \\ \mu_Y\end{pmatrix}, \begin{pmatrix}\Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY}\end{pmatrix}\right)$$

The marginal distribution $p(X) = \int p(X, Y)dY$

is given by $p(X) = N(\mu_X, \Sigma_{XX})$.

Formal definition

University of Basel

A Gaussian process $p(u) = GP(\mu, k)$

is a probability distribution over functions

$$u : \mathcal{X} \rightarrow \mathbb{R}^d$$

such that every finite restriction to function values

$$u_X = (u(x_1), \dots, u(x_n))$$

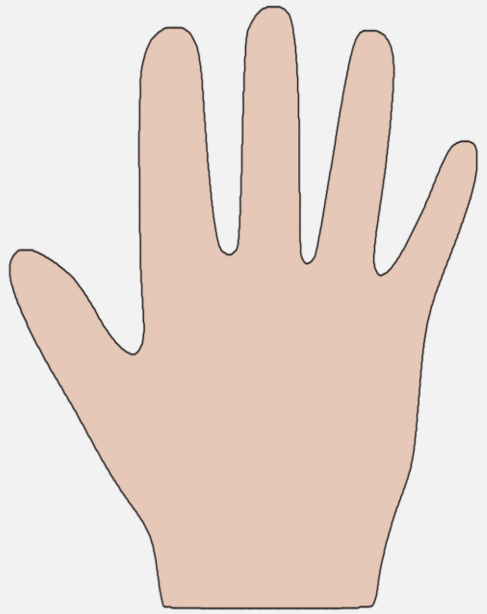
is a multivariate normal distribution

$$p(u_X) = N(\mu_X, k_{XX}).$$

Marginalization

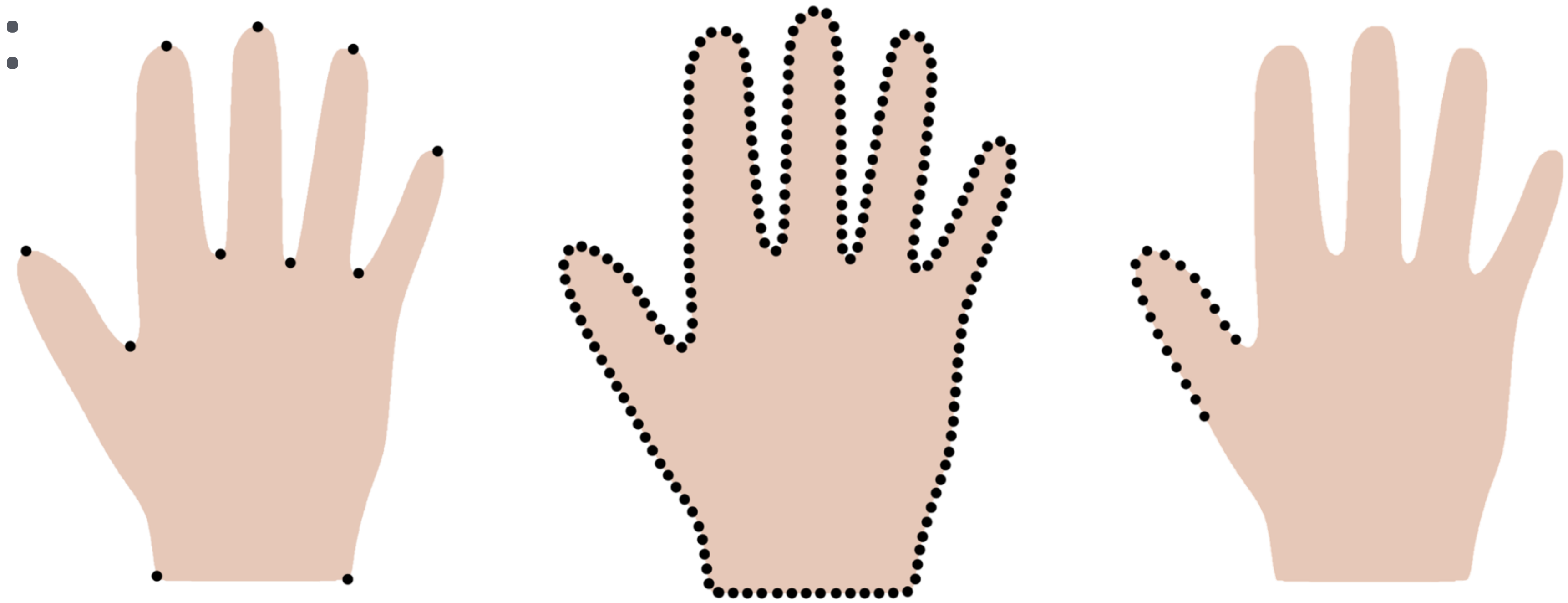
Continuous representation:

$$GP(\mu, k)$$

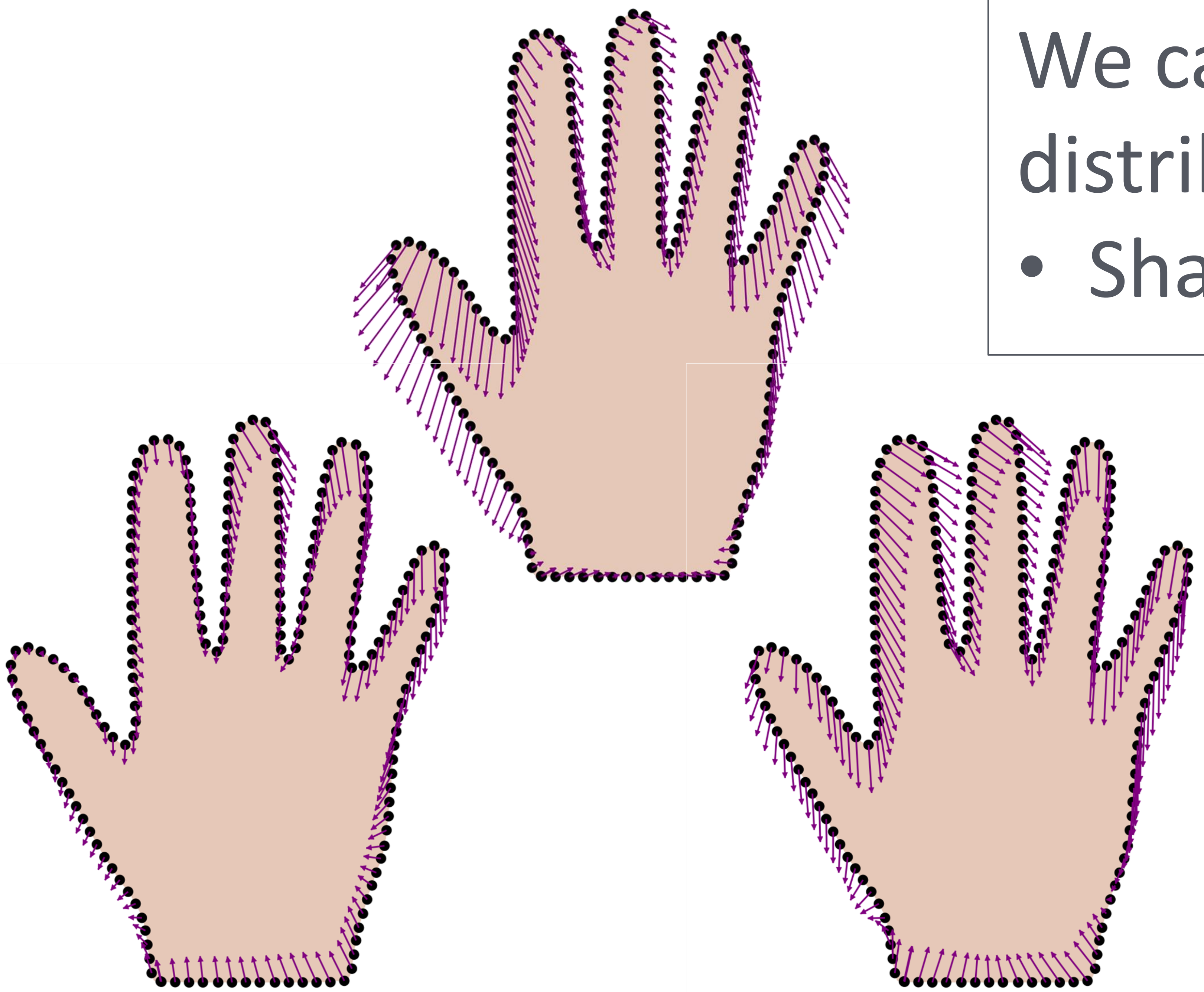


Discrete representation:

$$N(\mu, K)$$



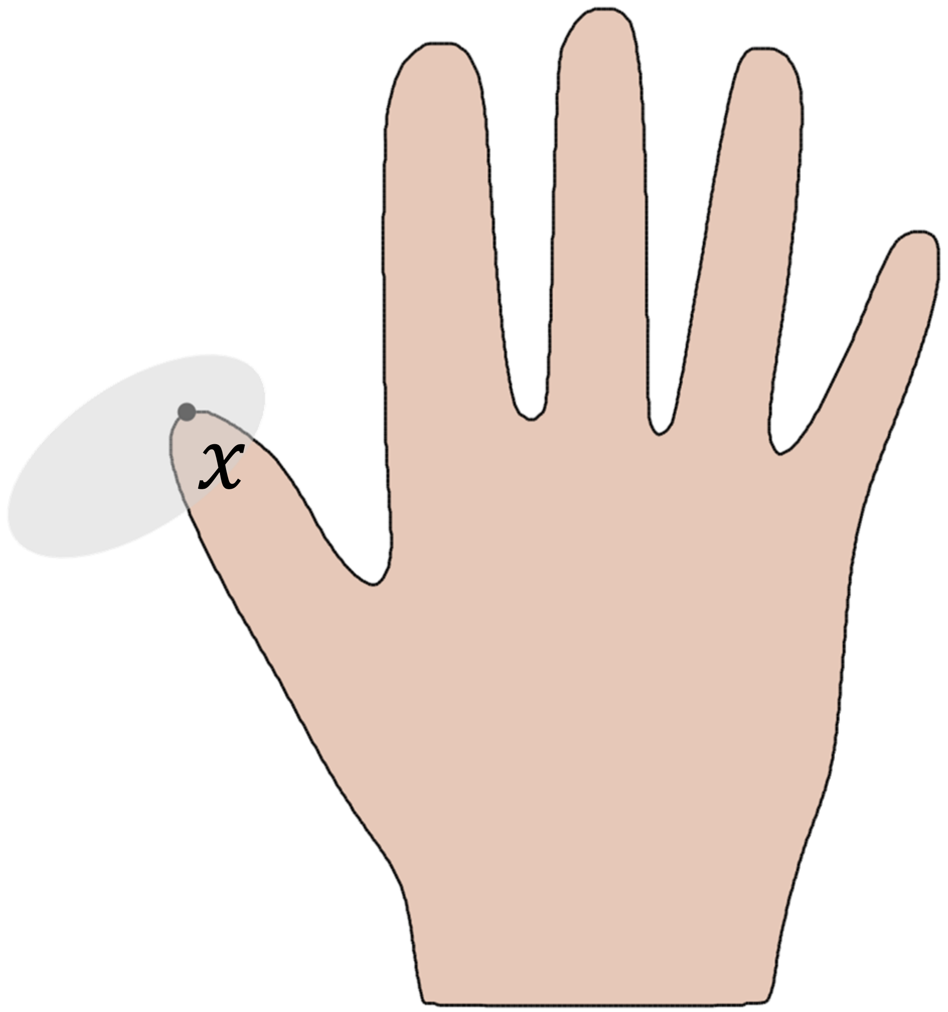
Sampling



We can sample from a normal distribution.

- Shape variations can be visualized

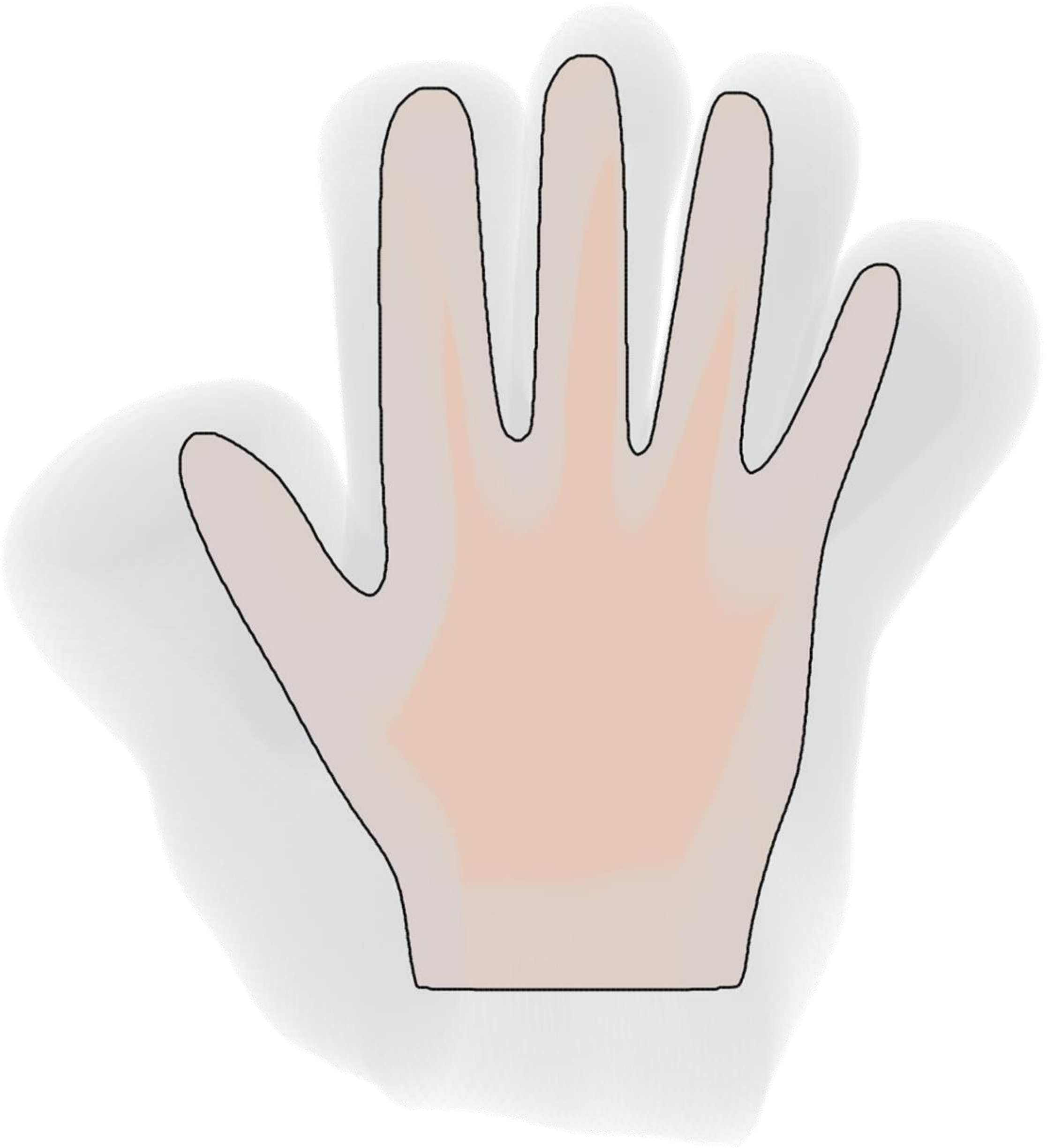
Computing probabilities



We can compute probabilities to answer questions like:

- What is the “normal” variation?
- Is a given variation still normal?

Computing probabilities



We can us the marginal distributions to compute confidence regions.