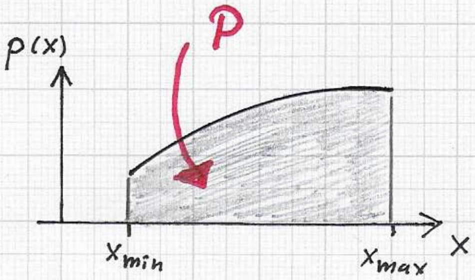
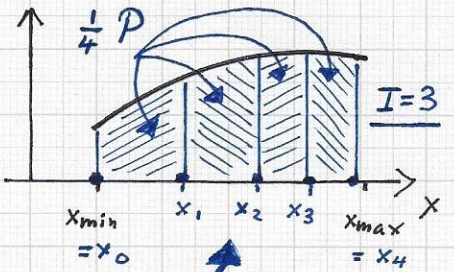
GEOMETRIC MODELING AND DENSITY – Cont'd.

- **Integrated Probability:** We can also use "integrated probability" to determine where to place 'representative feature points' in feature space, and how to establish these points' (local) density – given an initial large and high-dimensional feature point sample set. We first consider the continuous univariate **(1D)** case. Assuming





uniform distribution of integrated probability, i.e.:

$\int_{x_0}^{x_1} p(x)\,dx = \ldots = \int_{x_3}^{x_4} p(x)\,dx.$

that feature (point) values vary between x_{min} and x_{max}, an associated distribution or "probability" function $p(x)$ defines the relative occurrences of x-values. The goal is to subdivide the x-domain, i.e., $[x_{min}, x_{max}]$, into sub-intervals in such a way that **INTEGRATED PROBABILITY OVER EACH SUB-INTER-VAL IS THE SAME.**

The original integrated probability is given as

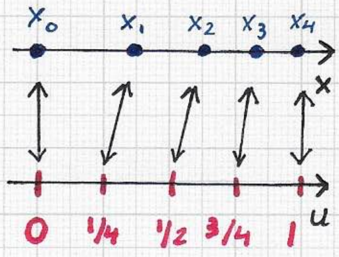
$$P = \int_{x_{min}}^{x_{max}} p(x)\,dx \ .$$

(It can be advantageous to normalize: $P = 1$.)
To subdivide $[x_{min}, x_{max}]$ into intervals $[x_i, x_{i+1}]$, $i = 0 \ldots I$, one must determine x_i-values as follows:

$$x_0 = x_{min} \ ; \ \int_{x_i}^{x_{i+1}} p(x)\,dx = P/(I+1), \ i = 0 \ldots I \ .$$

■ GEOMETRIC MODELING AND DENSITY - Cont'd.

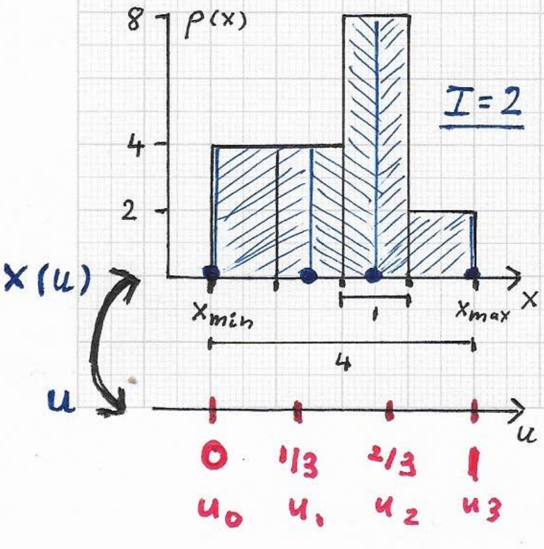
• **Integrated Probability:**



Mapping parameter space to feature space (u-space to x-space mapping)

Once the x_i-values are known, one can associate uniformly, equidistantly distributed parameter values u_i with them, where $u_i = i/(I+1)$, $i=0\ldots(I+1)$. The eventual goal is the computation of a low-degree (quadratic or piecewise quadratic) least-squares best approximation, subject to the mapping conditions $u_i \mapsto x_i$. The resulting best approximation serves as a model that – when evaluated at an arbitrary resolution of equidistant u-values – generates synthetic feature point distributions "of class $p(x)$."

➔ Since the probability function $p(x)$ is not known analytically and is only approximately known via a discrete, binned histogram representation, the method must be adapted to the discrete setting. Consider the example (left):



➔ $P = \int_{x_{min}}^{x_{max}} p(x)\,dx = 4+4+8+2 = 18$

➔ wanted: 4 values x_i such that $\int_{x_i}^{x_{i+1}} p(x)\,dx = 6$

➔ resulting x_i-values: $x_0 = x_{min}$, $x_1 = x_0 + \frac{3}{2}$, $x_2 = x_1 + 1$, $x_3 = x_2 + \frac{3}{2} = x_{max}$

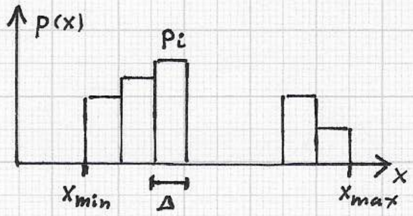
➔ associated u_i-values: $u_i = i/(I+1) = i/3$, $i = 0\ldots3$.

➔ The resulting model $x(u)$ generates "same-class histograms." ~BH

Stratovan

■ GEOMETRIC MODELING AND DENSITY – Cont'd.

• **Integrated Probability:**

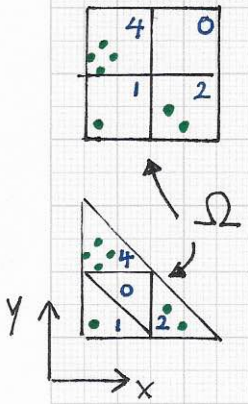


original histogram

The general 1D discrete case is based on on a histogram, defined for a domain interval $[x_{min}, x_{max}]$ divided into bins of constant width Δ with bin values p_i. Integrated probability is $\Delta \sum_i p_i = P$. To split the domain interval into sub-intervals with the same integrated probability $P/(I+1)$ the x_i-values are obtained by the conditions $\int_{x_i}^{x_{i+1}} p(x)\, dx = P/(I+1)$, $i = 0 \ldots I$. The associated parameter values are $u_i = i/(I+1)$, $i = 0 \ldots (I+1)$.

parametrization

• **2D Case:** In the bivariate case, feature points lie in the 2D xy-plane, defining a bivariate probability function $p(x,y)$. For example, a discrete representation of $p(x,y)$ could be represented via a tiling / binning of the xy-domain consisting of rectangles or triangles. First, one computes the integrated probability as $P = \iint_{\Omega} p(x,y)\, dx\, dy$, where the integration domain Ω is assumed to consist of one region with a boundary "tightly enclosing" the feature points. The goal is to subdivide the domain Ω into 2D sub-domains such that the integrated probability over each sub-domain is the same and the sum of all integrated probabilities over all sub-domains equals P.

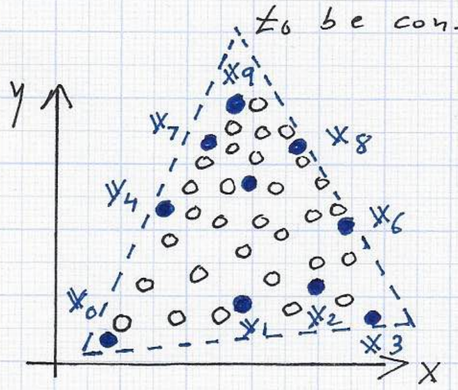


Tilings of xy-plane (feature space); feature points and their numbers are shown (per tile).

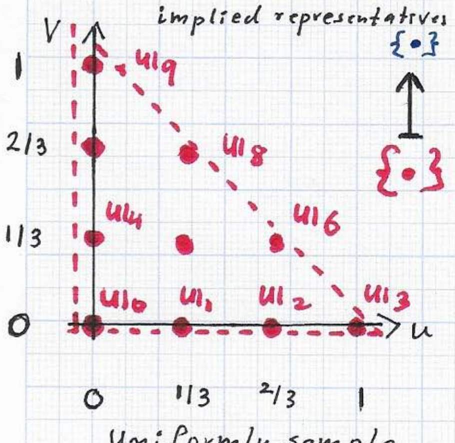
papersnake.com

■ GEOMETRIC MODELING AND DENSITY – Cont'd.

• **Algorithm:** We consider the 2D, bivariate setting to describe the processing steps. Given is a set of scattered feature points in a region in the plane. These points define the distribution that the model to be constructed must be able to replicate.



Given data {o} and implied representatives {•}



Uniformly sample parameter space

The example (left) shows the original feature point set {o} and the selected and/or constructed smaller data set {•}. The set {•} has been constructed in such a way that {o} and {•} have (nearly) the same (normalized) point probability distribution in the xy-plane. The goal is to compute a (quadratic) mapping that is obtained by associating a parameter tuple $\mathbf{u}_i = (u_i, v_i)$ of a set of uniformly, equidistantly placed parameter tuples {•} to a corresponding constucted point $\mathbf{x}_i = (x_i, y_i)^T$.

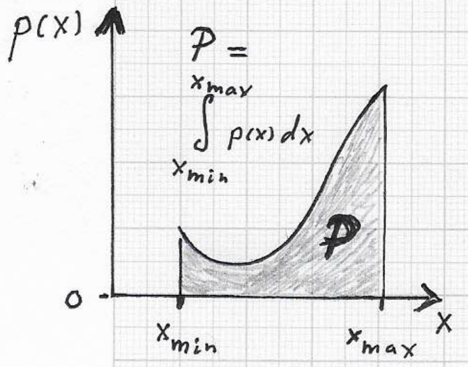
The best quadratic approximation obtained by mapping $\mathbf{u}_i$ to $\mathbf{x}_i$ has six coefficient $\mathbf{c}_{\mathbf{j}}$: $\mathbf{x}(\mathbf{u}) = \sum_{\|\mathbf{j}\| \le 2}' \mathbf{c}_{\mathbf{j}} \mathbf{u}^{\mathbf{j}} =$
$= \mathbf{c}_{0,0} + \mathbf{c}_{1,0}\, u + \mathbf{c}_{0,1}\, v + \mathbf{c}_{1,1}\, uv + \mathbf{c}_{2,0}\, u^2 + \mathbf{c}_{0,2}\, v^2$.

The coefficients $\mathbf{c}_{\mathbf{j}} = \mathbf{c}_{i,j}$ have an x- and y-coordinate component. The mapping $\mathbf{u}_i \mapsto \mathbf{x}_i$, $i = 0 \ldots 9$ (here), defines the overdetermined system $\sum_{\|\mathbf{j}\| \le 2} \mathbf{c}_{\mathbf{j}} (\mathbf{u}_i)^{\mathbf{j}} = \mathbf{x}_i$, solved via the least-squares method.

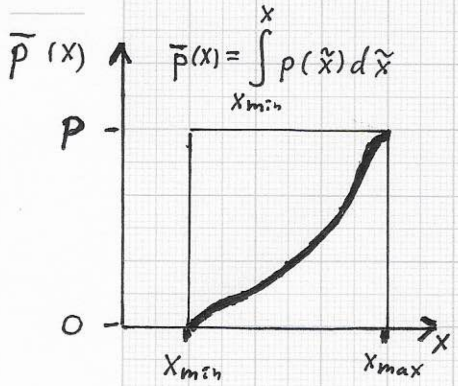
The model $\mathbf{x}(\mathbf{u})$ has an associated error.

GEOMETRIC MODELING AND DENSITY – Cont'd.

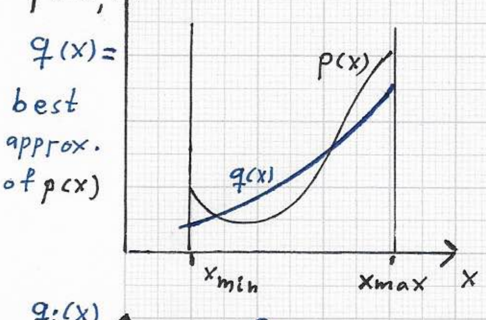
• Best approximation (analytical case):

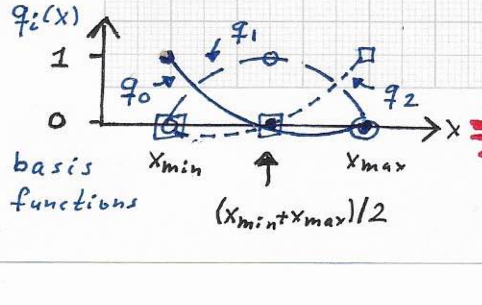


analytical form of $p(x)$



cumulative distribution





It is also possible to consider all feature points/vectors given in high-dimensional feature space for the computation of a (relatively) low-degree polynomial best approximation, capable of generating random, synthetic feature point distributions representing the same class. First, one must define and analytically represent (simply!) the given discrete feature point data distribution (e.g., using an efficient piecewise constant definition). Second, one applies best approximation methods (analytical/continuous setting) to compute a best quadratic approximation of the given distribution, for example. The illustrations depict the case of a 1D feature space: (i) $p(x)$ is the given distribution function, having the associated cumulative distribution $\bar{p}(x)$. (ii) $q(x)$ is the best quadratic approximation of $p(x)$. (iii) The quadratic basis functions $q_i(x)$, defined for the 3 knots $x_0 = x_{min}$, $x_1 = \frac{1}{2}(x_{min}+x_{max})$, $x_2 = x_{max}$. They satisfy the condition $q_i(x_j) = \delta_{i,j}$. Thus, the best approximation is $q(x) = \sum_{i=0}^{2} c_i\, q_i(x)$:

$$\begin{bmatrix} \langle q_0(x), q_0(x)\rangle & \cdots & \langle q_0(x), q_2(x)\rangle \\ \vdots & & \vdots \\ \langle q_2(x), q_0(x)\rangle & \cdots & \langle q_2(x), q_2(x)\rangle \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} \langle p(x), q_0(x)\rangle \\ \vdots \\ \langle p(x), q_2(x)\rangle \end{bmatrix}.$$