Part 2: Regression

$Y \in \mathbb{R}$, $X \in \mathbb{R}^D$ "true" model (unknown "data generating process"):

- deterministic: $y = f^*(x)$ $\quad f^*$: true generative process
- non-deterministic $y = f^*(x, \eta)$ ← random numbers

for fixed x, $f^*(x, \eta)$ will not always produce the same result, because η will be different in every run

⇒ determine the posterior prob. of all possible outcomes y:

$$p(y|x) \iff y = f^*(x, \eta) \text{ with } X \perp \eta$$

(f^*: deterministic fct. with random argum.)

"reparameterization trick" of $p(y|x)$

- simplest case: additive noise: $y = f^*(x, \eta) = f(x) + \eta$
  ⇒ task: find a model $\hat{f}(x) \approx f(x)$
- general case: outcome depends on the noise in a more complicated way
  example: post-additive noise: $y = g(f(x) + \eta)$
- nature of the function $f(x)$:
  - simplest case: linear $\quad f(x) = x \cdot \beta$ ← column vector of weights
    ↑ row vector
    "linear regression"
  - otherwise: "non-linear regression"
- noise η can always be uniform, because $\eta' = Q(\eta)$ can have arbitrary distr. with $Q(\eta)$ the "quantile function" absorbed into $f^*(x, \eta)$

Zoo of regression methods