Why

We want to estimate a random vector $x : \Omega \to \mathbf{R}^d$ from a random vector $y : \Omega \to \mathbf{R}^n$.

Definition

Denote by $g : \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}$ the joint density for (x, y).[1] Denote the conditional density for x given y by $g_{x|y} : \mathbf{R}^d \times \mathbf{R}^n \to \mathbf{R}$. In this setting, $g_{x|y}$ is called the *posterior density*, g_x is called the *prior density*, and $g_{y|x}$ is called the *likelihood density* and g_y is called the *marginal likelihood density*.

As usual (and assuming $g_y > 0$), the posterior is related to the likelihood, prior and marginal likelihood by

$$g_{x|y} \equiv \frac{g_x g_{y|x}}{g_y}.$$

A *maximum conditional estimate* for $x : \Omega \to \mathbf{R}^n$ given that y has taken the value $\gamma \in \mathbf{R}^n$ is a maximizer $\xi \in \mathbf{R}^d$ of $g_{x|y}(\xi, \gamma)$. It is also called the *maximum a posteriori estimate* or *MAP estimate*. The maximum conditional estimate is natural, in part, because it also maximizes the joint density, since $g(\xi, \gamma) = g_y(\gamma) g_{x|y}(\xi, \gamma)$ for all $\xi \in \mathbf{R}^d$ and $\gamma \in \mathbf{R}^n$.

---

[1]Future editions will comment on the existence of such a density.