(1)

$$\int \mathcal{L}(f, y, x)\, \underbrace{p(x,y)}_{\text{data dist}^{n}.}\, dx\, dy$$

$(x_1, y_1) \ldots \quad (x_n, y_n)$

$\sim p(x,y).$

$$\approx \frac{1}{m} \sum_{i=1}^{m} \mathcal{L}(f, y_i, x_i)$$

$y_i \quad \tilde{y}_i$

standard way $\longrightarrow \mathcal{L}(f, \underline{\tilde{y}_i}, x_i)$

$P(y_i | \tilde{y}_i)$

true y. counts

$$\frac{1}{n} \sum_{i=1}^{m} \int \mathcal{L}(f, y_i, x_i)\, \underbrace{P(y_i|\tilde{y}_i)}_{\text{Dirichlet. posterior.}}\, dy_i$$

w/samples.

sgd.

$\approx$

---

(2)

$$\min_{\theta} \int \mathcal{L}(f_\theta, y, x)\, \underline{p(y|x)}\, \underline{p(x)}\, dx\, dy$$

$\downarrow$

0/1

$p(y|x)$ is continuous.

Dirichlet for estimating $p(y|x)$

Sampling from posterior to capture uncertainty.