Notations. All along these notes, we will use the following notations for the variables:

$Y =$ observed variables;

$Z =$ unobserved (hidden, latent) variables;

$\theta =$ parameters;

$x =$ covariates (if needed).

As for the distributions, we will denote

$f(\cdot)$ **or** $p(\cdot) =$ probability distribution function (pdf);

$f_\theta(\cdot) = f(\cdot;\theta)$ **or** $p_\theta(\cdot) = p(\cdot;\theta) =$ pdf with parameter θ;

$\mathbb{E}_\theta =$ expectation under p_θ.

The subscript θ may be replaced by the distribution itself (e.g. $\mathbb{E}_p$ or $\mathbb{E}_q$) or dropped when not necessary.

As for classical distributions, we will use the following notations:

$\mathcal{U}_{[a,b]} =$ uniform distribution over the interval $[a, b]$;

$\mathcal{N}(\mu, \sigma^2) =$ Gaussian distribution with mean μ and variance σ^2;

$\mathcal{M}(n, \pi) =$ multinomial distribution with n draws and vector of probabilities $\pi = (\pi_1, \dots, \pi_K)$, $(\sum_k \pi_k = 1)$;

$\mathcal{P}(\lambda) =$ Poisson distribution with mean λ;

$\mathbf{B}(\alpha, \beta) =$ beta distribution;

$\mathcal{D}(\alpha) =$ Dirichlet distribution with parameter $\alpha = (\alpha_1, \dots, \alpha_K)$;

$\mathcal{NB}(\pi, r) =$ negative binomial distribution with probability π and number of successes r;

$\mathcal{G}\mathbf{am}(a, b) =$ gamma distribution with shape parameter a and rate b.

We will also use the abbreviations *rv* for 'random variable', *iid* for 'independent and identically distributed' and *wrt* for 'with respect to'. We will denote $[\![i, j]\!] = \{i, i+1, \dots, j-1, j\}$.