These notes summarise selected lecture concepts and are not a substitute for working through the lecture slides, tutorials, and self-study exercises. Feel free to personalise and develop them into your own summary sheet.

Loss $\ell(\mathbf{h}, \mathbf{a})$ — We take action $\mathbf{a}$ without observing $\mathbf{h}$; once $\mathbf{h}$ is revealed, we incur loss $\ell(\mathbf{h}, \mathbf{a})$.

Risk $R(\mathbf{a}; f)$ — The expected loss $\mathbb{E}_{f(\mathbf{h})}[\ell(\mathbf{h}, \mathbf{a})]$, with f denotes the distribution over $\mathbf{h}$, potentially being a prior belief $p(\mathbf{h})$, a posterior $p(\mathbf{h}|\mathbf{x})$, or the data distribution.

Empirical risk $\hat{R}_n(\mathbf{a}; f)$ — The risk when the expectation over f is approximated with a sample average based on n samples $\mathbf{h}_i \sim f(\mathbf{h})$, $\hat{R}_n(\mathbf{a}; f) = \frac{1}{n}\sum_{i=1}^{n} \ell(\mathbf{h}_i, \mathbf{a})$.

Expected loss principle — We choose the action $\mathbf{a}^*(f)$ that minimises the (empirical) risk.

Policy — A mapping from evidence to action.

Common loss functions and actions that minimise the risk—

loss	$\ell(\mathbf{h}, \mathbf{a})$	optimal action $\mathbf{a}^*(f)$
quadratic	$(h-a)^2$	$\mathbb{E}_{f(h)}[h]$
absolute error	$\lvert h-a \rvert$	median(f)
0-1	$1 - \mathbb{1}(\mathbf{h} = \mathbf{a})$	$\text{argmax}_{\mathbf{h}} \log f(\mathbf{h})$
log$^{(*)}$	$-\log q(\mathbf{h})$	f

$^{(*)}$ For the log-loss, the action $\mathbf{a}$ is a distribution $q(\mathbf{h})$.

Maximum likelihood estimation — Frequentist view: minimising empirical risk under log-loss. Bayesian view: minimising the posterior expected 0-1 loss with a flat prior.