(1) b) d)

$$C(g) = -\frac{2}{n} \sum g(X_i) + \int g(x)^2 dx$$

$$C(g,f) = -\frac{2}{n} \sum \int g(x) g_0(x)\, dx + \int g(x)^2\, dx =$$

$$= -2 \int g(x) g_0(x)\, dx + \int g(x)^2\, dx = \int (g(x) - f(x))^2 dx - \int g(x)^2 dx =$$

$$= \int (g(x) - f(x))^2\, dx - \int g(x)^2\, dx =$$

$$= \int (\cancel{(g(x))^2} - 2 f(x) g(x) + (f(x))^2 - \cancel{g(x)^2}\, dx) =$$

$$= \int (-2 g(x) f(x) + f(x)^2)\, dx = \int (g(x) - f(x))^2_{dx} - \int f(x)^2) dx$$

this is minimal when $\int g(x) dx = \int f(x) dx$

(1) c) Let $\hat{p}_h$ denote the kernel estimator on bandwidth h. Assume the sample size is evend and denote it by $2n$. Randomly split the data $X = (X_1 .. X_{2n})$ into two sets of size n. Denote these by $y = (y_1 .. y_n)$ and $z = (z_1 .. z_n)$. Let $H = \{h_1 .. h_N\}$ be a finite grid of the bandwidths. Let

$$\hat{p}_j(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{h_j^d} K\left(\frac{\|x - y_i\|}{h}\right)$$

. Thus we have a set $P = \{\hat{p}_1 .. \hat{p}\}$ of density estimators.

We would like to minimize $L(p, \hat{p}_j) = \int \hat{p}_j^2(x) - 2 \int \hat{p}_j(x) p(x)\, dx$.

Define the estimated risk

$$\hat{L}_j \equiv \hat{L}(p, \hat{p}_j) = \int \hat{p}_j^2(x) - \frac{2}{n} \sum_{i=1}^{n} \hat{p}_j(z_i)$$

Let $\hat{p} = \operatorname{argmin}_{g \in P} \hat{L}(p, g)$.

$$X = (X_1 .. X_{2n}) \overset{\text{split}}{\Rightarrow} \begin{array}{l} y \to \{\hat{p}_1, .. \hat{p}_N\} = P \\ z \to \{\hat{L}_1 .. \hat{L}_N\} \end{array}$$