Optimal Estimation

Consider the following series of measurements of the unknown value x:

$$y_i = a_i x + \varepsilon_i, \quad i = 1, \dots, n,$$

where y_i are measurement results, a_i are known coefficients, and ε_i represent random error of measurement and are independent identically distributed (i.i.d.) with zero mean and variance σ^2:

$$\mathsf{E}\varepsilon_i = 0, \quad \mathsf{E}\varepsilon_i^2 = \sigma^2, \quad i = 1, \dots, n,$$

a) What function of $y_1, \dots, y_n$ (and of $a_1, \dots, a_n$) would you use as a good estimate $\widehat{x}$ for x? $\widehat{x} =$? (Hint: try best linear estimate or least squares estimate).

b) Is this estimate optimal in any sense?

c) Is it a biased or an unbiased estimate?

d) What is its variance (expressed through σ^2)? $\mathrm{Var}(\widehat{x}) =$?

e) How would you estimate σ^2 if it is unknown? $\widehat{\sigma^2} =$?

f) What would you use as an estimate for $\mathrm{Var}(\widehat{x})$ if σ^2 is unknown? $\widehat{\mathrm{Var}(\widehat{x})} =$?

g) Suppose that the variance σ^2 is known. What "canonical information" would be sufficient to extract from the series of records

$$(y_1, a_1), \dots, (y_n, a_n), \quad i = 1, \dots, n$$

in order to compute the estimate $\widehat{x}$, and its variance $\mathrm{Var}(\widehat{x})$?

h) Suppose that the variance σ^2 is NOT known. What "canonical information" would be sufficient to extract from the series of observations in order to compute $\widehat{x}$, $\widehat{\sigma^2}$, and $\widehat{\mathrm{Var}(\widehat{x})}$?

i) How should we update such "information" when a new record (y_{n+1}, a_{n+1}) arrives?

j) How should we "combine" (merge) two pieces of "canonical information"?

Please do not try to use general formulas, but develop as much as possible from scratch.