1.8.a. Let the random variables y_i be IID with Bernoulli distribution with probability p of success, $i = 1, \ldots, n$. So $P[y_i = 1] = p$ and $P[y_i = 0] = 1-p$, and this discrete probability distribution can be written as $f_p(y_i) = p^{y_i}(1-p)^{1-y_i}$, $y_i = 0, 1$. The likelihood and log-likelihood are

$$L(p) = f_p(y_1, \ldots, y_n) = \prod_{i=1}^{n} \{p^{y_i}(1-p)^{1-y_i}\} = p^{\sum y_i}(1-p)^{n-\sum y_i},$$
$$\log L(p) = \sum y_i \log(p) + (n - \sum y_i) \log(1-p).$$

The first order derivative of the log-likelihood is

$$\frac{d \log L(p)}{dp} = \frac{\sum y_i}{p} - \frac{n - \sum y_i}{1-p}.$$

The maximum of this function is obtained by solving

$$\frac{\sum y_i}{p} - \frac{n - \sum y_i}{1-p} = 0.$$

This gives $(1-p)\sum y_i - p(n - \sum y_i) = 0$ with solution $\hat{p}_{\mathrm{ML}} = \frac{1}{n}\sum y_i = \bar{y}$. This is a maximum, as the second order derivative is

$$\frac{d^2 \log L(p)}{dp^2} = -\frac{\sum y_i}{p^2} - \frac{n - \sum y_i}{(1-p)^2} < 0.$$

The mean and variance of $\hat{p}_{\mathrm{ML}}$ are

$$E[\hat{p}_{\mathrm{ML}}] = E\left[\frac{\sum y_i}{n}\right] = \frac{1}{n}\sum E[y_i] = p, \qquad \text{(S 1.5)}$$
$$\mathsf{var}\,(\hat{p}_{\mathrm{ML}}) = \frac{1}{n^2}\mathsf{var}\left(\sum y_i\right) = \frac{1}{n^2}\sum \mathsf{var}\,(y_i) = \frac{n}{n^2}p(1-p) = \frac{p(1-p)}{n}. \qquad \text{(S 1.6)}$$

From (S 1.5) we see that $\hat{p}_{\mathrm{ML}}$ is unbiased, and from (S 1.6) that for $n \to \infty$ the variance tends to zero. Therefore the conditions in (1.48) are satisfied so that $\hat{p}_{\mathrm{ML}}$ is a consistent estimator.

b. The Cramér-Rao lower bound for the variance is the inverse of the 1×1 information matrix

$$\mathcal{I} = -E\left[\frac{d^2 \log L(p)}{dp^2}\right] = -E\left[-\frac{\sum y_i}{p^2} - \frac{n - \sum y_i}{(1-p)^2}\right]$$
$$= \frac{np}{p^2} + \frac{n - np}{(1-p)^2} = \frac{n}{p} + \frac{n}{1-p} = \frac{n}{p(1-p)}.$$

The inverse is

$$\mathcal{I}^{-1} = \frac{p(1-p)}{n}. \qquad \text{(S 1.7)}$$

c. Equation (S 1.5) shows that $E[\hat{p}_{\mathrm{ML}}] = p$, and (S 1.7) proves that the variance of $\hat{p}_{ML}$ in (S 1.6) is minimal, therefore the estimator is efficient (in the class of all unbiased estimators of p).

d. Let $r = \hat{p}/(1-\hat{p})$ denote the estimated odds ratio. If $0 < p \leq 1$ then there is a positive probability of p^n that all outcomes are $y_i = 1$, $i = 1, \cdots, n$, in which case $\hat{p} = 1$. This means that r takes non-negative values and, with probability p^n, an unbounded value. Therefore $E[r] = \infty$ so that r is not unbiased.

If $0 \leq p < 1$ then r is a consistent estimator of the odds ratio. This can be proved by the result of part (c) that $\mathrm{plim}(\hat{p}) = p$, so that $\mathrm{plim}(1-\hat{p}) = 1 - \mathrm{plim}(\hat{p}) = 1 - p \neq 0$ and

$$\mathrm{plim}\left(\frac{\hat{p}}{1-\hat{p}}\right) = \frac{\mathrm{plim}(\hat{p})}{\mathrm{plim}(1-\hat{p})} = \frac{p}{1-p}.$$