Deviations from the mean, entropy typical set

Submit a PDF of your answers to Canvas.

1. This problem we will explore bounding the probability that the sample mean deviates from the true mean for a Bernoulli random variables. Consider a sequence of independent coin flips, denoted $X_1, X_2, \ldots, X_n$. The coin flips are i.i.d. Bernoulli random variables:

$$X_i = \begin{cases} 0 & \text{with probability } 1-\theta \\ 1 & \text{with probability } \theta \end{cases}$$

You flip the coin n times and observe k heads (where heads corresponds to $X_i = 1$).

a) Find/state the expression for $\widehat{\theta}_{\mathrm{ML}}$, i.e., the maximum likelihood estimate of θ.

b) Find an exact expression for the probability that $\widehat{\theta}_{\mathrm{ML}}$ deviates from the true mean by more than $\delta \in [0, 1]$:

$$\mathbb{P}(|\theta - \widehat{\theta}_{\mathrm{ML}}| \geq \delta)$$

Recall the Binomial pmf, which gives the probability of k heads in n coin flips:

$$\binom{n}{k}\theta^k(1-\theta)^{n-k}$$

c) Use Hoeffding's inequality to derive a bound on the probability that $\widehat{\theta}_{\mathrm{ML}}$ deviates from the true mean by more than δ.

d) Use a computer to evaluate the exact expression and Hoeffding's inequality when $\theta = 0.4$, $n = 10000$, and $\delta = 0.1$.

e) *Typically*, one would expect that the number of heads (after n flips) would be about θn. One precise notion of *typical* is based on the Shannon information. Recall that if the outcome of a sequence of coin flips is $x_1, \ldots, x_n$, the Shannon information (in bits) is given by $\log_2(1/p(x_1, \ldots, x_n))$. If X_i are independent, the Shannon information is $\sum_i \log_2(1/p(x_i))$. Define the set of outcomes that have approximately average Shannon information as the *entropy* typical set:

$$\left\{(x_1, \ldots, x_n) : \left|\frac{1}{n}\log_2\left(\frac{1}{p(x_1, \ldots, x_n)}\right) - H(X_i)\right| \leq \epsilon\right\}.$$

Find an upper bound on the probability that an outcome is not in the entropy typical set by using Hoeffding's inequality.