

Problem Set 4

Spring 2018

Issued: February 9, 2018

Due: 11:59 PM, Monday, February 12, 2018

1. Transform Practice

Consider a random variable Z with transform

$$M_Z(s) = \frac{a - 3s}{s^2 - 6s + 8}, \quad \text{for } |s| < 2.$$

Calculate the following quantities:

- (a) The numerical value of the parameter a .
- (b) $\mathbb{E}[Z]$.
- (c) $\text{var}(Z)$.

2. Bounds for the Coupon Collector's Problem

Recall the coupon collector's problem, where X is a random variable which is equal to the number of boxes bought until one of every type of coupon is obtained (there are n total coupons).

The expected value of X is nH_n , where H_n is the *harmonic number of order n* which is defined as

$$H_n \triangleq \sum_{i=1}^n \frac{1}{i},$$

and satisfies the inequalities

$$\ln n \leq H_n \leq \ln n + 1.$$

- (a) Use Markov's inequality in order to show that

$$\mathbb{P}(X > 2nH_n) \leq \frac{1}{2}.$$

- (b) Use Chebyshev's inequality in order to show that

$$\mathbb{P}(X > 2nH_n) \leq \frac{\pi^2}{6(\ln n)^2}.$$

Note: You can use the identity

$$\sum_{i=1}^{\infty} \frac{1}{i^2} = \frac{\pi^2}{6}.$$

(c) Define appropriate events and use the union bound in order to show that

$$\mathbb{P}(X > 2nH_n) \leq \frac{1}{n}.$$

Note: The sequence $a_n = (1 - 1/n)^n$, for $n = 1, 2, 3, \dots$, is strictly increasing and $\lim_{n \rightarrow \infty} a_n = 1/e$.

3. A Chernoff Bound for the Sum of Coin Flips

Let X_1, \dots, X_n be i.i.d. Bernoulli(q) random variables with bias $q \in (0, 1)$, and call X their sum, $X = X_1 + \dots + X_n$, which a Binomial(n, q) random variable, with mean $\mathbb{E}[X] = nq$.

(a) Let $\epsilon > 0$ such that $q + \epsilon < 1$, and define $p = q + \epsilon$. Show that for any $t > 0$,

$$\mathbb{P}(X \geq pn) \leq \exp(-n(tp - \ln \mathbb{E}[e^{tX_1}])).$$

(b) The *Kullback-Leibler divergence* from the distribution Bernoulli(q) to the distribution Bernoulli(p), is defined as

$$D(p \parallel q) \triangleq p \ln \frac{p}{q} + (1 - p) \ln \frac{1 - p}{1 - q}.$$

The Kullback-Leibler divergence can be interpreted as a measure of how close the two distributions are. One motivation for this interpretation is that the Kullback-Leibler divergence is always nonnegative, i.e. $D(p \parallel q) \geq 0$, and $D(p \parallel q) = 0$ if and only if $p = q$. So it can be thought of as a ‘distance’ between the two Bernoulli distributions.

Optimize the previous bound over $t > 0$ and deduce that

$$\mathbb{P}(X \geq pn) \leq e^{-nD(p \parallel q)}.$$

(c) Moreover, the Kullback-Leibler divergence is related to the square distance between the parameters p and q via the following inequality

$$D(p \parallel q) \geq 2(p - q)^2, \quad \text{for } p, q \in (0, 1).$$

Use this inequality in order to deduce that

$$\mathbb{P}(X \geq (q + \epsilon)n) \leq e^{-2n\epsilon^2},$$

and

$$\mathbb{P}(X \leq (q - \epsilon)n) \leq e^{-2n\epsilon^2}.$$

Hint: For the second bound use symmetry in order to avoid doing all the work again.

(d) Conclude that

$$\mathbb{P}(|X - qn| \geq \epsilon n) \leq 2e^{-2n\epsilon^2}.$$

4. Decoding a Bit from a Noisy Signal

In many wireless communications systems, each receiver listens on a specific frequency. The bit b sent is represented by a $+1$ or -1 . Unfortunately, noise from other nearby frequencies can affect the receiver's signal. A simplified model for this noise is as follows. There are n other senders. The i th sender is also trying to send a bit B_i that is represented by $+1$ or -1 . The receiver obtains the signal S given by

$$S = b + w \sum_{i=1}^n B_i,$$

where w is constant indicating the power of the bits of the other senders.

In order to decode b from S , we use the following scheme: if S is closer to $+1$ than -1 , the receiver assumes that the bit sent was a $+1$; if S is closer to -1 than $+1$, the receiver assumes that the bit sent was a -1 ; if S is equidistant to $+1$ and -1 , the receiver fails to recover b .

Assume that all the bits B_i are independent and uniformly distributed over $\{+1, -1\}$.

- Show that the probability that the receiver cannot determine b correctly, is at most $2 \exp(-\frac{1}{2nw^2})$.
Hint: Transform appropriately each B_i in order to use Problem 3.
- If we want to ensure that the probability to correctly determine b is at least $1 - \delta = 0.999$, what condition do we need to impose on the power of the noise w ?
- What would be the condition on the power of the noise w , if we have used Chebyshev's inequality in order to upper bound the error probability?
- Discuss how the analysis of the error probability in (a) compares with the analysis of the error probability using Chebyshev's inequality.

5. [Bonus] Gaussian Tail Bounds

The bonus question is just for fun. You are not required to submit the bonus question, but do give it a try and write down your progress.

Let $\phi(y) = \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}}$ be the PDF of a standard normal random variable $Y \sim \mathcal{N}(0, 1)$.

- Show that for $y \neq 0$ we have that

$$\phi(y) = -\frac{1}{y} \cdot \phi'(y).$$

- Use (a) to show that

$$\mathbb{P}(Y \geq t) \leq \frac{1}{t} \cdot \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}}, \quad \text{for all } t > 0.$$

(c) Use part (a) to show that

$$\left(\frac{1}{t} - \frac{1}{t^3}\right) \frac{e^{-\frac{t^2}{2}}}{\sqrt{2\pi}} \leq \mathbb{P}(Y \geq t), \quad \text{for all } t > 0.$$