PHYS 7895 Fall 2015
Introduction to Quantum Information Theory
Homework 1

Due Friday 27 January 2017, by 3pm in Nicholson 447
(Please try all the problems by yourself first. If you find that you struggle with them, you are allowed to work with no more than one collaborator as long as you write down who your collaborator is. No late assignments will be accepted. Please be sure to download the latest version of the notes before starting the homework.)

This assignment has a first part and a second part.

First part: Exercises in **http://www.markwilde.com/qit-notes.pdf**:

2.1.1, 2.2.1

Second part: The following exercises:

1. Concentration inequalities:

   (a) Prove the Markov inequality. That is, for a random variable X whose realizations are non-negative, prove that

   $$\Pr\{X \geq \varepsilon\} \leq \frac{\mathbb{E}\{X\}}{\varepsilon}.$$

   (b) Prove the Chebyshev inequality. That is, for any random variable with finite second moment, show that the following inequality holds:

   $$\Pr\{|X - \mathbb{E}\{X\}| \geq \varepsilon\} \leq \frac{\mathrm{Var}\{X\}}{\varepsilon^2},$$

   where $\mathrm{Var}\{X\} = \mathbb{E}\{|X - \mathbb{E}\{X\}|^2\}$.

   (c) Prove the following law of large numbers. For a large number of pairwise independent and identically distributed random variables $X_1, \ldots, X_n$, (such that $\mathbb{E}\{X_i\} = \mu$ and $\mathbb{E}\{|X_i - \mu|^2\} = \sigma^2$ for all $i \in \{1, \ldots, n\}$) the probability that the sample mean deviates from the true mean has a power law decay:

   $$\Pr\left\{\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq \varepsilon\right\} \leq \frac{\sigma^2}{\varepsilon^2 n}.$$

   (d) Prove the first part of the Chernoff-Hoeffding bound. That is, for a large number of bounded independent and identically distributed random variables $X_1, \ldots, X_n$, show that the probability that the sample mean deviates from the true mean by an additive constant (one-sided) decays exponentially with the number of samples taken:

   $$\Pr\left\{\frac{1}{n}\sum_{i=1}^{n} X_i - \mu \geq \varepsilon\right\} \leq \inf_{t>0} \frac{[\mathbb{E}_X\{\exp\{tX\}\}]^n}{\exp\{t(\mu + \varepsilon)\}^n}.$$

The idea to finish it off from here is to choose t small enough so that we have

$$[\mathbb{E}_X\{\exp\{tX\}\}]/\exp\{t(\mu+\varepsilon)\} < 1$$

(implying an exponential decay with n). Bonus points for taking it from here to get the exponential decay.

2. Before you were born, the mathematicians (probabilists) were hard at work, trying to obtain ever finer characterizations of the convergence rate in the central limit theorem. The initial best characterizations are due to Berry and Esseen, who proved the following theorem. Let $Z_1, \ldots, Z_n$ be a sequence of i.i.d. random variables (assume finite cardinality for simplicity and each with mean μ and variance σ^2). Then the deviation of the tail of the sample mean of the normalized $Z_1, \ldots, Z_n$ from the tail of a normalized Gaussian falls off as the inverse square-root of the number of samples:

$$\left| \Pr\left\{ \frac{\sum_{i=1}^{n}[Z_i - \mu]}{\sqrt{n\sigma^2}} \geq \delta_1 \right\} - Q(\delta_1) \right| \leq \frac{C\xi}{\sigma^3\sqrt{n}},$$

where ξ is the third central moment of each Z_i, C is a fixed positive constant that seems to keep improving, and $Q(x)$ is the tail of a standard Gaussian:

$$Q(x) \equiv \frac{1}{\sqrt{2\pi}} \int_x^\infty \exp\left\{-|u|^2/2\right\}\, du.$$

Let us define the set $\mathfrak{T}^{X^n}$ of "one-sided" variance-typical sequences for a distribution $p_X(x)$ to be as follows:

$$\mathfrak{T}^{X^n} \equiv \left\{ x^n : -\frac{1}{n}\log(p_{X^n}(x^n)) - H(X) \leq \sqrt{\frac{V(X)}{n}} Q^{-1}(\varepsilon) \right\},$$

where $V(X) = \mathrm{Var}_X\{-\log p_X(X)\}$, ε is a fixed positive constant, and Q^{-1} is the inverse of the Q function. (The above definition differs from the one we defined in class just by focusing on one side of the typicality tolerance and by choosing the typicality parameter δ to decrease with n, i.e., $\delta = \sqrt{V(X)/n}\, Q^{-1}(\varepsilon)$.)

(a) Find an upper bound on the size of the set $\mathfrak{T}^{X^n}$.

(b) Using the Berry-Esseen theorem, find an upper bound on the probability that a random sequence X^n falls outside the set $\mathfrak{T}^{X^n}$.

(c) Put these two facts together and describe a data compression scheme (Shannon-style). That is, for a specified error probability ε' (and such that n is of the order $(1/\varepsilon')^2$), to how many bits can this scheme allow for us to compress a random length-n sequence?[1]

[1] The Berry-Esseen theorem has been put to great use recently in many problems in both classical and quantum information theory to account for finite-size effects that get "washed away" in the limit of many instances of a resource.