On Distinguishability Measures for Quantum States

Christopher Granade

August 9, 2010

Outline

But First...

...An Apology.

My voice is weak today, so my slides are more verbose than they probably should be in order to make up. Thanks for understanding.

What is a State?

The way we think about what a quantum state is will give us some hints as to how to think about distinguishing states.

Operationalist View

We can think of a quantum state as a calculational tool to generate probability distributions for hypothetical measurements. In particular, for a state ρ, if we wish to measure some property X represented by a POVM $\{M_x\}_{x \in X}$, we obtain that:

$$p(x) = \Pr(X = x) = \mathrm{tr}(M_x \rho)$$

Classical Distributions

The operationalist view suggests that we can think of state distinguishability in terms of *classical* distributions.

Classical Problem

How distinguishable are two probability distributions $p_0(x)$ and $p_1(x)$ over the same random variable X?

We will often put this problem in terms of a random variable $T = \{ 0, 1 \}$ that picks one of the distributions.
We suppose that $\Pr(T = 0) = \Pr(T = 1) = 1/2$. Thus, $p_t(x) = \Pr(X = x | T = t)$ so that the distribution over X is given by:

$$p(x) = \sum_{t \in T} \Pr(T = t, X = x) = \frac{1}{2} \sum_{t \in T} p_t(x)$$

Cryptographic Distinguishability Measures

Fuchs and de Graaf (1998) consider four particular measures of distinguishability useful in cryptography:

- Probability of error.
- Kolmogorov distance.
- Bhattacharyya coefficient.
- Shannon distinguishability.

Each of these can then be generalized to a measure of state distinguishability by optimizing over measurements.

Definition

The probability of error $\mathrm{PE}(p_0, p_1)$ is the total probability of incorrectly guessing which distribution was used to generate a sample x. Here, the optimal strategy is to always pick the distribution most likely to have produced x.

$$\begin{aligned}\mathrm{PE}(p_0, p_1) &= \sum_{x \in X, t \in T} \Pr(X = x, T = t) \Pr(\text{error} | X = x, T = t) \\ &= \frac{1}{2} \sum_{x \in X} \min(p_0(x), p_1(x))\end{aligned}$$

Application to States

Let $\mathrm{PE}(\rho_0, \rho_1)$ be the minimum over all POVMs of the classical probability of error. It is known that the optimal measurement gives the explicit form:

$$\mathrm{PE}(\rho_0, \rho_1) = \frac{1}{2} + \frac{1}{2} \sum_{\lambda_j \leq 0} \lambda_j$$

where λ_j are the eigenvalues of $\rho_0 - \rho_1$.
We can therefore relate the probability of error to the *trace norm*:

$$\begin{aligned} \mathrm{PE}(\rho_0, \rho_1) &= \frac{1}{2} + \frac{1}{4} \sum_j (\lambda_j - |\lambda_j|) \\ &= \frac{1}{2} + \frac{1}{4} \cancelto{0}{\mathrm{tr}(\rho_0 - \rho_1)} - \frac{1}{4} \mathrm{tr}\,|\rho_0 - \rho_1| \end{aligned}$$

Definition

If we think of the probability distribution functions p_0 and p_1 as vectors, then the Kolmogorov distance between them is half of the L1 norm of their difference:

$$\mathrm{K}(p_0, p_1) = \frac{1}{2} \sum_{x \in X} |p_0(x) - p_1(x)|$$

A little algebra yields that:

$$\begin{aligned} \mathrm{K}(p_0, p_1) &= 1 - 2\,\mathrm{PE}(p_0, p_1) \\ \mathrm{K}(\rho_0, \rho_1) &= 1 - 2\left(\frac{1}{2} - \frac{1}{4}\,\mathrm{tr}\,|\rho_0 - \rho_1|\right) \\ &= \frac{1}{2}\,\mathrm{tr}\,|\rho_0 - \rho_1| \end{aligned}$$

Definition and Optimization

Whereas the Kolmogorov distance can be thought of as the L1 norm between two vectors, the Bhattacharyya coefficient is a natural inner product on the space of probability distributions:

$$\mathrm{B}(p_0, p_1) = \sum_{x \in X} \sqrt{p_0(x) p_1(x)}$$

By explicitly optimizing, Fuchs and Caves (1998) demonstrated that:

$$\mathrm{B}(\rho_0, \rho_1) = \min_{\mathcal{M}} \mathrm{B}(\mathcal{M}(\rho_0), \mathcal{M}(\rho_1)) = \mathrm{tr}\left(\sqrt{\sqrt{\rho_0}\rho_1\sqrt{\rho_0}} \right)$$

where $\mathcal{M}$ is a POVM which induces a distribution $\mathcal{M}(\rho)$.

Relation to Fidelity

Recall that the fidelity between two pure states is given by their inner product:

$$F(|\psi\rangle, |\phi\rangle) = |\langle\psi \,|\, \phi\rangle|$$

By Uhlmann's theorem,

$$\mathrm{B}(\rho_0, \rho_1) = \max_{|\psi_0\rangle, |\psi_1\rangle} F(|\psi_0\rangle, |\psi_1\rangle)$$

where the maximization is taken over purifications of ρ_0, ρ_1. Thus, we see that the Bhattacharyya coefficient tells us how much two states overlap. Fully overlapping states are completely indistinguishable.

Motivation and Definition

The final measure that we consider is motivated by considering the uncertainty involved in distinguishing two distributions. In particular, when we sample the rv X, our uncertainty about whether our sample was drawn from p_0 or p_1 may be reduced:

$$\begin{aligned} \mathrm{SD}(p_0, p_1) &= \text{uncertainty before sampling} - \text{after sampling} \\ &= H(T) - H(T|X) = I(T|X) \end{aligned}$$

Since $I(T|X) = I(X|T)$, we can directly calculate the Shannon distinguishability:

$$\mathrm{SD}(p_0, p_1) = H(p) - \frac{1}{2}\left(H(p_0) + H(p_1)\right)$$

Note that $\mathrm{SD}(\rho_0, \rho_1)$ has no closed form, due to ln being transcendental.

Bounding B and K

It is rather inconvenient to have four measures of the same thing. Each of these measures has advantages and disadvantages, and so we would like to know how they relate. We do so by deriving inequalities which bound the various measures.
In the interests of time, we shall focus on the B and K measures, and shall show that:

$$1 - \mathrm{B}(p_0, p_1) \leq \mathrm{K}(p_0, p_1) \leq \sqrt{1 - \mathrm{B}^2(p_0, p_1)}$$

Proof (1/2)

We start by showing $1 - \mathrm{B}(p_0, p_1) \leq \mathrm{K}(p_0, p_1)$. To do so, we utilize that $\sum(p_0(x) + p_1(x)) = 2$, so that we can factor B.

$$\begin{aligned}
1 - \mathrm{B}(p_0, p_1) &= \frac{1}{2} \sum_{x \in X} \left[p_0(x) + p_1(x) - 2\sqrt{p_0(x) p_1(x)} \right] \\
&= \frac{1}{2} \sum_{x \in X} (\sqrt{p_0(x)} - \sqrt{p_1(x)})^2 \\
&= \frac{1}{2} \sum_{x \in X} \left| \sqrt{p_0(x)} - \sqrt{p_1(x)} \right|^2 \\
&\leq \frac{1}{2} \sum_{x \in X} |p_0(x) - p_1(x)| = \mathrm{K}(p_0, p_1)
\end{aligned}$$

Proof (2/2)

$$
\begin{aligned}
\mathrm{K}^2(p_0, p_1) &= \frac{1}{4}\left(\sum_{x\in X}|p_0(x) - p_1(x)|\right)^2 \\
&= \frac{1}{4}\left(\sum_{x\in X}\left|\sqrt{p_0(x)} - \sqrt{p_1(x)}\right|\left|\sqrt{p_0(x)} + \sqrt{p_1(x)}\right|\right)^2 \\
&\quad \text{(via Schwarz ineq)} \\
&\leq \frac{1}{4}\left[\sum_{x\in X}(\sqrt{p_0(x)} - \sqrt{p_1(x)})^2\right]\left[\sum_{x\in X}(\sqrt{p_0(x)} + \sqrt{p_1(x)})^2\right] \\
&= \frac{1}{4}(2 - 2B(p_0, p_1))(2 + 2B(p_0, p_1)) = 1 - B(p_0, p_1)
\end{aligned}
$$

$$
\begin{aligned}
\sum_{x\in X}(\sqrt{p_0(x)} \pm \sqrt{p_1(x)})^2 &= \sum_{x\in X}(p_0(x) + p_1(x) \pm \sqrt{p_0(x)p_1(x)}) \\
&= 2 \pm 2\,\mathrm{B}(p_0, p_1)
\end{aligned}
$$

Monotonicity of Optimization (1/2)

We next show that these inequalities continue to hold when we consider the state distinguishability measures B and K.
Let $\mathcal{E}_{\mathsf{B}}$ be a POVM optimizing B. Define K similarly.
Then, for any POVM $\mathcal{E}'$, since $\mathcal{E}_{\mathsf{K}}$ maximizes K:

$$\mathsf{K}(\mathcal{E}'(\rho_0), \mathcal{E}'(\rho_1)) \leq \mathsf{K}(\mathcal{E}_{\mathsf{K}}(\rho_0), \mathcal{E}_{\mathsf{K}}(\rho_1))$$

In particular:

$$1 - \mathsf{B}(\rho_0, \rho_1) \leq \mathsf{K}(\mathcal{E}_{\mathsf{B}}(\rho_0), \mathcal{E}_{\mathsf{B}}(\rho_1)) \leq \mathsf{K}(\mathcal{E}_{\mathsf{K}}(\rho_0), \mathcal{E}_{\mathsf{K}}(\rho_1))$$

Thus, $1 - \mathsf{B}(\rho_0, \rho_1) \leq \mathsf{K}(\rho_0, \rho_1)$.

Monotonicity of Optimization (1/2)

Since B is minimized rather than maximized:

$$\mathsf{B}(\mathcal{E}_{\mathsf{B}}(\rho_0), \mathcal{E}_{\mathsf{B}}(\rho_1)) \leq \mathsf{B}(\mathcal{E}'(\rho_0), \mathcal{E}'(\rho_1))$$

The rest of the derivation follows similarly:

$$\begin{aligned}
\mathsf{K}(\rho_0, \rho_1) &= \mathsf{K}(\mathcal{E}_{\mathsf{K}}(\rho_0), \mathcal{E}_{\mathsf{K}}(\rho_1)) \\
&\leq \sqrt{1 - \mathsf{B}^2(\mathcal{E}_{\mathsf{K}}(\rho_0), \mathcal{E}_{\mathsf{K}}(\rho_1))} \\
&\leq \sqrt{1 - \mathsf{B}^2(\mathcal{E}_{\mathsf{B}}(\rho_0), \mathcal{E}_{\mathsf{B}}(\rho_1))} \\
&= \sqrt{1 - \mathsf{B}^2(\rho_0, \rho_1)}
\end{aligned}$$

Asymptotics

In security-related applications (such as QKD analysis), it is common to speak of two *sequences* of density matrices which must be indistinguishable in the asymptotic limit. We say that such sequences are *exponentially indistinguishable*.
For instance, with respect to K, $\{\rho_{n,0}\}$ and $\{\rho_{n,1}\}$ are EI if there exists ϵ such that:

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : K(\rho_{n,0}, \rho_{n,1}) \leq \epsilon^n$$

One of the key results of Fuchs and de Graaf is that EI with respect to any of PE, K, B, SD implies EI with respect to all.

All four measures considered here have benefits and disadvantages. Inequalities relate these measures and guarantee that for exponential indistinguishablity, it doesn't matter which you use.