36710-36752, Fall 2020
Homework 2

Due Mon, Oct 5 by 5pm.

1. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $(S, \mathcal{A})$ a measurable space and $f\colon \Omega \to S$ a measurable function.. Show that, for arbitrary subsets $A, A_1, A_2, \ldots$ of S,

   (a) $f^{-1}(A^c) = \left(f^{-1}(A)\right)^c$;

   (b) $f^{-1}\left(\cup_n A_n\right) = \bigcup_n f^{-1}(A_n)$ and

   (c) $f^{-1}\left(\cap_n A_n\right) = \bigcap_n f^{-1}(A_n)$.

   (The last two identities actually hold also for uncountable unions and intersections). Let $\mathcal{A}$ be a σ-field over S. Prove that the collection $f^{-1}(\mathcal{A}) = \{f^{-1}(A), A \in \mathcal{A}\}$ of subsets of Ω is a σ-field over Ω (in fact, the smallest σ-field on Ω that makes f measurable).

2. (**The induced measure is a measure**) Let $(\Omega, \mathcal{F}, \mu)$ be a measure space, $(S, \mathcal{A})$ a measurable space and $f\colon \Omega \to S$ a measurable function. Show that the measure induced by f and μ, i.e. the function ν over $\mathcal{A}$ given by

$$A \mapsto \mu\left(f^{-1}(A)\right), \quad A \in \mathcal{A},$$

   is a measure. Show by example that ν need not be σ-finite if μ is σ-finite.

3. Let f be an integrable real-valued function over a measure space $(\Omega, \mathcal{F}, \mu)$. Show that f is finite almost everywhere $[\mu]$. *Hint: you may assume that* $f \geq 0$ *(why?); then, you only need to show that* $f < \infty$ *almost everywhere* $[\mu]$.

4. Let $\{f_n\}_n$ be a sequence of real valued function on some measurable space $(\Omega, \mathcal{F})$. Show that the set $\{\omega\colon\ \lim_n f_n(\omega) \text{ exists}\}$ is measurable. *Hint: express* $\lim_n f_n(\omega)$ *using* $\text{limsup}_n f_n(\omega)$ *and* $\text{liminf}_n f_n(\omega)$*, which are measurable...*

5. Let P and Q two probability measures on some measurable space $(\Omega, \mathcal{F})$ and let μ be any σ-finite measure on that space such that both P and Q are absolutely continuous with respect to μ (for example, you may take $\mu = P + Q$). Let $p = \frac{dP}{d\mu}$ and $q = \frac{dQ}{d\mu}$ be the corresponding Radon-Nikodym derivatives.
   The total variation distance between P and Q is defined as

$$d_{\text{TV}}(P, Q) = \sup_{A \in \mathcal{F}} |P(A) - Q(A)|.$$

   This is a very strong notion of distance between probability distributions: if $d_{\text{TV}}(P, Q) < \epsilon$, for some small $\epsilon > 0$, then the P-probability and Q-probability of *any* event will differ by at most ϵ. It can be showed that d_{TV} is a metric over the set of all probability measure on $(\Omega, \mathcal{F})$. In particular, $d_{\text{TV}}(P, Q) = 0$ if and only if $P = Q$.

(a) Show that $d_{\text{TV}}(P, Q) = 1$ if and only if P and Q are mutually singular.

(b) Prove that the following equivalent representation of the total variation distance:

$$d_{\text{TV}}(P, Q) = \frac{1}{2} \int_{\Omega} |p - q| d\mu.$$

Thus, the total variation distance is half the L_1 distance between densities.
Hint: in the definition of total variation distance you may want to take $A = \{q \geq p\}$. *Show that the supremum is achieved by this set...*

(c) **Total variation distance and hypothesis testing.** Let X be a random variable taking values in some measurable space $(S, \mathcal{A})$. Suppose we are interested in testing the null hypothesis that the distribution of X (a probability measure on $(S, \mathcal{A})$!) is P versus the alternative hypothesis that it is Q. We do so by devising a *test* ϕ, which is a measurable function from S into $\{0, 1\}$ such that $\phi(x) = 1$ (resp. $\phi(x) = 0$) signifies that the null hypothesis is rejected (resp. not rejected) if X takes on the value x. To measure the performance of a given test function ϕ we evaluate its risk, defined as the sum of type I and type II errors:

$$R_{P,Q}(\phi) = \int_S \phi dP + \int_S (1 - \phi) dQ.$$

Show that

$$\inf_{\phi} R_{P,Q}(\phi) = 1 - d_{\text{TV}}(P, Q),$$

where the infimum si over all test functions.
The above result formalizes the intuition that the closer P and Q are, the harder it is to tell them apart using any test function. In particular, $R_{P,Q}(\phi) = 0$ – i.e., it is possible to perfectly discriminate between P and Q – if and only if the two probability measures are mutually singular.
Hint: use the Neymann-Pearson approach and take ϕ *to be the indicator function of the set* $\{q \geq p\}$.