Statistics 330b/600b, Math 330b spring 2017
Homework # 7
Due: Thursday 9 March

*[1] For real-valued random variables X_1 and X_2, define

$$f(\omega, s) := \mathbb{1}\{X_1(\omega) > s\} + \mathbb{1}\{X_2(\omega) > s\} - 2\mathbb{1}\{X_1(\omega) > s, X_2(\omega) > s\}.$$

Show that $\int_{\mathbb{R}} f(\omega, s)\, ds = |X_1(\omega) - X_2(\omega)|$. Then complete the argument begun in class to show that $\mathbb{P}|X_1 - X_2|$ is minimized over all integrable random variables with $X_i \sim P_i$ if the X_i's are coupled using the quantile transformation.

*[2] Suppose $\mathcal{A}$ is a sigma-field on a set $\mathcal{X}$ and $\mathcal{B}$ is a countably generated sigma-field on a set $\mathcal{Y}$, that is, $\mathcal{B} = \sigma(\mathcal{E})$ for some countable $\mathcal{E} \subseteq \mathcal{B}$. Suppose also that $\mathcal{B}$ separates the points of $\mathcal{Y}$: if $y_1 \neq y_2$ then there exists a set $B \in \mathcal{B}$ for which $y_1 \in B$ and $y_2 \in B^c$. Without loss of generality $\mathcal{E}$ is stable under the formation of complements.

Suppose T is an $\mathcal{A}\backslash\mathcal{B}$-measurable map from $\mathcal{X}$ into $\mathcal{Y}$. Define $\text{graph}(T) := \{(x, Tx) : x \in \mathcal{X}\}$, a subset of $\mathcal{X} \times \mathcal{Y}$.

(i) For $y_1 \neq y_2$, explain why there exists a set $E \in \mathcal{E}$ for which $\mathbb{1}_E(y_1) \neq \mathbb{1}_E(y_2)$.

(ii) Define $H := \cup_{E \in \mathcal{E}} (T^{-1}(E^c)) \times E$. Show that $H \subseteq \text{graph}(T)^c$.

(iii) If $y \neq Tx$, with $x \in \mathcal{X}$ and $y \in \mathcal{Y}$, show that $(x, y) \in H$.

(iv) Deduce that $\text{graph}(T)^c = H$ and hence $\text{graph}(T) \in \mathcal{A} \otimes \mathcal{B}$.

*[3] Let $\{X_i : i \in \mathbb{N}\}$ be a set of random variables all defined on the same $(\Omega, \mathcal{F}, \mathbb{P})$. Define $X(\omega) = (X_1(\omega), X_2(\omega), \dots)$, which you know can be thought of as an $\mathcal{F}\backslash\mathcal{B}$-measurable map from Ω into $\mathbb{R}^{\mathbb{N}}$, for the product sigma-field $\mathcal{B} = \mathcal{B}(\mathbb{R})^{\mathbb{N}}$. Let P denote the distribution of X.

Suppose f in $\mathcal{L}^1(\mathbb{R}^{\mathbb{N}}, \mathcal{B}, P)$. Show that, for each $\epsilon > 0$, there exists a $k \in \mathbb{N}$ and a Borel measurable function g_k on $\mathbb{R}^k$ for which $\mathbb{P}|f(X) - g_k(X_1, \dots, X_k)| < \epsilon$.

[4] For each fixed $p > 0$ define $B_{p,n} := \{x \in \mathbb{R}^n : g_p(x) \leq 1\}$ where $g_p(x) := \sum_{i \leq n} |x_i|^p$. Follow these steps to show that

$$V_{p,n} := \text{vol}(B_{p,n}) = \frac{(2\Gamma(1 + 1/p))^n}{\Gamma(1 + n/p)}.$$

Remark. Recall that the Gamma function is defined for $\alpha > 0$ by $\Gamma(\alpha) := \int_0^\infty t^{\alpha - 1} e^{-t} dt$. Integration by parts shows that $\Gamma(1 + \alpha) = \alpha\Gamma(\alpha)$ for each $\alpha > 0$.

(i) For $x \in \mathbb{R}^n$ Show that

$$I_p := \int_{\mathbb{R}^n} \exp\left(-g_p(x)\right)\, dx = \left(2 \int_0^\infty \exp(-t^p)\, dt\right)^n = \left(\frac{2}{p}\Gamma(1/p)\right)^n.$$

(ii) Show that

$$\exp\left(-g_p(x)\right) = \int_0^\infty \{t \geq g_p(x)\} e^t\, dt = \int_0^\infty \{t^{-1/p} x \in B_{p,n}\} e^{-t}\, dt$$

(iii) Deduce that $I_p = V_p \int_0^\infty t^{n/p} e^{-t} dt$.

(iv) Then what?

[5]

(i) Let P be a probability measure on $\mathcal{B}(\mathbb{R})$. Define

$$m_0 := \inf\{x : P(-\infty, x] \geq 1/2\}.$$

Show that $P[m_0, \infty) \geq 1/2$ and $P(-\infty, m_0] \geq 1/2$. *[Such a value m_0 is called a median for P.]*

(ii) Suppose $Z = X + Y$, with X and Y independent random variables. Let m be a median for the distribution of Y. Show that $\mathbb{P}\{X \geq x\} \leq 2\mathbb{P}\{Z \geq x + m\}$ for each real x.