Problem Set 1

Data Analysis and Interpretation (EE 223)

Instructor: Prof. Prasanna Chaporkar
EE Department, IIT Bombay

1. Let $\{x_1, x_2, \ldots, x_n\}$ be a given data set.
   (a) Find $c^\star \in \arg\min_{c \in \Re} \sum_{i=1}^n (x_i - c)^2$.
   (b) Find $m^\star \in \arg\min_{m \in \Re} \sum_{i=1}^n |x_i - c|$.

2. Define $\tilde{\mathcal{S}}_k = \{i : x_i - \overline{x} \geqslant ks\}$. Show that
$$\frac{\tilde{\mathcal{S}}_k}{n} \leqslant \frac{1}{1+k^2} \quad \text{for every } k \geqslant 1.$$

3. Let $\mathcal{X} = \{x_1, \ldots, x_n\}$ and $\mathcal{Y} = \{y_1, \ldots, y_n\}$ be two data sets, where x_i and y_i be two different parameter values corresponding to the index i. Define sample correlation between the two data sets as
$$r = \frac{\sum_{i=1}^n (x_i - \overline{x})(y_i - \overline{y})}{(n-1)s_x s_y},$$
where $\overline{x}$ and $\overline{y}$ are sample means and s_x and s_y are the sample standard deviations of the data sets $\mathcal{X}$ and $\mathcal{Y}$, respectively. Show that
   (a) $-1 \leqslant r \leqslant 1$
   (b) if $y_i = a + bx_i$ for every $i = 1, \ldots, n$ and $b \neq 0$, then $r = sign(b)$.

4. Show that power set of the set of natural numbers has the same cardinality as the set of real numbers.

5. Let $\{\mathcal{F}_i\}_{i \in I}$ be a collection of σ-fields on Ω. Show that $\cap_{i \in I} \mathcal{F}_i$ is also a σ-fields on Ω.

6. Let $\mathcal{B}$ denote the Borel σ-fields on $\Re$. Then show the following:
   (a) $(a, b) \in \mathcal{B}$ for all reals $a < b$.
   (b) $[a, b) \in \mathcal{B}$ for all reals $a < b$.
   (c) $(a, b] \in \mathcal{B}$ for all reals $a < b$.
   (d) $[a, b] \in \mathcal{B}$ for all reals $a \leqslant b$.
   (e) set of all irrationals is in $\mathcal{B}$.

7. Let $A_1, A_2, \ldots$ be a sequence of events in measurable space $(\Omega, \mathcal{F})$. Define the following two sets
$$\overline{A} = \limsup_{n \uparrow \infty} A_n = \cap_{n=1}^\infty \cup_{k=n}^\infty A_k, \quad \text{and}$$
$$\underline{A} = \liminf_{n \uparrow \infty} A_n = \cup_{n=1}^\infty \cap_{k=n}^\infty A_k,$$
Show that
   (a) $\underline{A} \subseteq \overline{A}$.
   (b) $\underline{A}, \overline{A}$ are events.
   (c) $P(\overline{A}) \geqslant \limsup_{n \uparrow \infty} P(A_n)$.
   (d) $P(\underline{A}) \leqslant \liminf_{n \uparrow \infty} P(A_n)$.