Home Work 1: CAP 6610 Spring ’13
Due Date: Feb 8th 2013
Show all steps. Be as concise as possible.

1. Assume that we have two identical boxes: Box 1 and Box 2. Box 1 contains 5 red balls and 3 blue balls; Box 2 contains 2 red balls and 4 blue balls. A box is selected at random and exactly one ball is drawn from that box.

   – What is the probability that the ball is blue? (show all steps)
   – Given that the selected ball is blue, what is the probability that it came from box 2? (show all steps)

2. Let the random variable $X$ be representable as a sum of random variables $X = \sum_{i=1}^{n} X_i$. Show that, if $E[X_iX_j] = E[X_i]E[X_j]$ for every pair $i, j$ with $1 \leq i < j \leq n$, then $Var[X] = \sum_{i=1}^{n} Var[X_i]$.

3. Let $f(x)$ be a convex function and $X \in \{x_i : i = 1, 2, \cdots, N\}$ be a random variable with probabilities $P(x_i)$ where $\sum_i P(x_i) = 1$. Then prove that $f(E(X)) \leq E(f(X))$.

4. – For any nonnegative random variable $X$, prove that $E(X) = \int_0^{\infty} P(X > s) \, ds$.

   – Assume $\xi_i$ for $i = 1, 2, \cdots, n$ are i.i.d real-valued random variables with mean 0 and $E(\xi_i^2) = \sigma^2$. Also assume that all are bounded within range $[-C, C]$ by some positive constant $C$, with probability one, and $s \geq 0$. Then prove that

   \[
P(\sum_{i=1}^{n} \xi_i \geq s) \leq \exp \left( -\frac{n\sigma^2}{C^2} \psi \left( \frac{sC}{n\sigma^2} \right) \right)
\]

   where $\psi(z) = (1 + z) \log(1 + z) - z$ for $z \geq 0$.

   (Hint: First use Hoeffdings’ inequality, then use exp series, then the inequality for exp(x), and then minimize w.r.t $\lambda$)

5. (Programming) Consider the following probability density function on $(x, y)$:

   \[f(x) = \text{Uniform}[-2, 2] \text{ and } f(y|x) = \mathcal{N}(\mu, \sigma^2) \text{ where } \mu = 0.3x^3 - 0.6x^2 + 0.05x - 3 \text{ and } \sigma = 0.25.\]

   Generate 20 i.i.d samples from this distribution.

   Now find the minimum empirical square error estimator from the class of all 3rd degree polynomials. Do the same for the class of all 5th degree polynomials.

   Plot the data superimposed with the two polynomials. Report the polynomials and the error in both cases.