Problem Set #2: Random variables

1. Consider the random variable $X \sim U[-1,1]$. Derive the CDF and (for continuous cases) the density function for the following random variables.

   (a) $Y = X$
   (b) $Y = X^2$
   (c) $Y = \begin{cases} 0 & \text{if } X < 0 \\ 1 & \text{otherwise} \end{cases}$
   (d) $Y = X + 2$
   (e) $Y = X^3$
   (f) $Y = |X|$
   (g) $Y = \log(X+2)$
   (h) $Y = F(X)$, where $F(X)$ denotes the CDF of X
   (i) $Y = \begin{cases} 0 & \text{if } X \in [-1/2, 1/2] \\ X & \text{otherwise} \end{cases}$
   (j) $Z = F(Y)$, where $F(Y)$ is the CDF of Y as defined in the previous problem.

2. What is the expectation of each of the random variables in problem 1?

3. What is the variance for each of the random variables in problem 1?

4. Assume that $X_1, \ldots, X_n \sim i.i.d.\ N(\mu, \sigma^2)$. Assume σ is known. Which of the following are random variables? Please answer yes or no and explain briefly.

   (a) X_i, $i = 1, \ldots, n$.
   (b) $E(X_i)$, $i = 1, \ldots, n$.
   (c) $f(X_i) = \log(X_i)$
   (d) $E(f(X_i))$
   (e) $g(X_1, X_2) = \log(X_1 + X_2)$
   (f) $E(g(X_1, X_2)|X_2 = 2)$
   (g) $E(g(X_1, X_2)|X_2)$
   (h) $\bar{X}_n \equiv \frac{1}{n}\sum_{i=1}^n X_i$
   (i) $\sqrt{n}\left(\bar{X}_n - 3\right)$
   (j) $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^n X_i$
   (k) $|\bar{X}_n - 3|$
   (l) $\text{Prob}\left(|\bar{X}_n - 3| > \epsilon\right)$, for some small $\epsilon > 0$.
   (m) $\lim_{n\to\infty} |\bar{X}_n - 3|$

(n) $\frac{\bar{X}_n - 3}{\sigma}$

(o) $\frac{1}{\sigma}\phi\left(\frac{\bar{X}_n}{\sigma}\right)$, where $\phi(\cdots)$ denotes the standard normal density function.

(p) $\Phi\left(\frac{\bar{X}_n}{\sigma}\right)$, where $\Phi(\cdots)$ denotes the standard normal cumulative distribution function.

(q) $\text{Prob}\left(|\frac{\bar{X}_n}{\sigma}| > 1.96|\mu = 0\right)$

(r) $\mathbf{1}(X_i > 0)$, where $\mathbf{1}([\cdots])$ denotes the "indicator" function, and equals 1 when the event $[\cdots]$ is true, and 0 otherwise.

(s) $\frac{1}{n}\sum_{i=1}^{n}\mathbf{1}(X_i > 0)$

(t) $E\mathbf{1}(X_i > 0)$

(u) $\tilde{X}_i \equiv X_i \cdot \mathbf{1}(X_i > 0)$

(v) $E(\tilde{X}_i)$

(w) $\max\{X_1, \ldots, X_n\}$

(x) $T_n \equiv \mathbf{1}\left(|\frac{\bar{X}_n}{\sigma}| > 1.96\right)$.

(y) $\rho(\mu) \equiv \text{Prob}(T_n = 1|\mu)$.

5. **(Joint, conditional, and marginal distributions)** Consider the random variables X, Y, which are joint uniformly distributed on the unit square (i.e., the density $f(x, y) = 1$ on $x, y \in [0, 1]$). What is:

(a) $E(X)$
(b) $E(Y)$
(c) $E(X|Y)$
(d) $E(X|Y < \frac{1}{2})$
(e) $E(X|Y > X)$
(f) $E(X|Y > \frac{1}{2} + X)$
(g) $\text{Cov}(X, Y)$
(h) $\text{Cov}(X, Y|Y < X)$
(i) $\text{Corr}(X, Y)$
(j) $\text{Corr}(X, Y|Y < X)$

6. Consider two random variables X, Y. Suppose we know the marginal CDF's $F_X(x)$, $F_Y(y)$. Define x_s such that $F_X(x_s) = s$, and y_t such that $F_Y(y_t) = t$. We wish to make inference regarding the joint CDF $F(x_s, y_t)$.

(a) Can you derive a lower bound on $F(x_s, y_t)$?

(b) Can you derive an upper bound on $F(x_s, y_t)$?