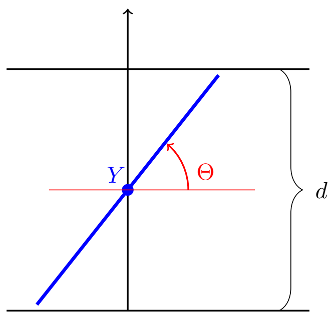
Middle East Technical University
Department of Electrical and Electronics Engineering
EE 230 Probability and Random Variables

Homework 4
Due 23:59 Saturday June 18, 2022

1. 

A length ℓ needle is dropped between two parallel lines separated by distance d in such a way that the distance of needle's midpoint from the first line, i.e., the random variable Y, and the angle needle makes with the lines, i.e., the random variable Θ, have the following joint probability density function

$$f_{Y,\Theta}(y,\theta) = \begin{cases} \frac{\cos\theta}{2d}, & \text{if } y \in [0,d], \text{ and } \theta \in [-\frac{\pi}{2}, \frac{\pi}{2}] \ , \\ 0, & \text{otherwise}, \end{cases} ,$$

where $d < \ell < 2d$. The needle may cross $y = 0$ and $y = d$ lines and the random variable N is the total number of lines crossed by the needle.

(a) What is the PMF p_N of N?

(b) What is the conditional PDF $f_{Y|\{N=2\}}$ of Y given $N = 2$?

(c) What is the conditional PDF $f_{\Theta|N}(\cdot|0)$ of Θ given $N = 0$?

2. (A midterm exam question from previous years) Consider random variables X and Y which are related as follows :

$$Y = g(X) = \begin{cases} 1, & -2 \le X < -1, \\ |X|, & -1 \le X < +1, \\ (X-2)^2, & +1 \le X \le +2 \\ 0. & \text{otherwise.} \end{cases}$$

a) If X is uniformly distributed in $[0, 1]$, determine the PDF $f_Y(y)$ explicitly and <u>plot</u> it.

b) If X is uniformly distributed in $[-1, 2]$, determine the PDF $f_Y(y)$ explicitly and <u>plot</u> it.

c) If X is uniformly distributed in $[-2, 2]$,

i. determine the CDF $F_Y(y)$ explicitly and <u>plot</u> it.

ii. what type of random variable is Y ?

d) (independent from previous parts) Z and W are independent continuous random variables, each uniformly distributed in $[0, 1]$. A random variable V is defined as $V = \dfrac{Z}{W+1}$. Find the PDF of V.

3. Parts a, b and c of this question are independent.

a) For each figure below (showing a rectangle, circle, ellipse, ellipse with primary axis rotated counter-clockwise by $\frac{\pi}{4}$ around its center), assume random variables X and Y are uniformly distributed in the shaded area and determine whether X and Y are independent and/or uncorrelated. If correlated, indicate whether the covariance is positive or negative. Clearly justify your answers.

b) The random variable X is uniform in $[0, 5]$. The random variable Y is defined as

$$Y = \begin{cases} 3X + 1, & X < 2 \\ X, & \textit{otherwise} \end{cases}$$

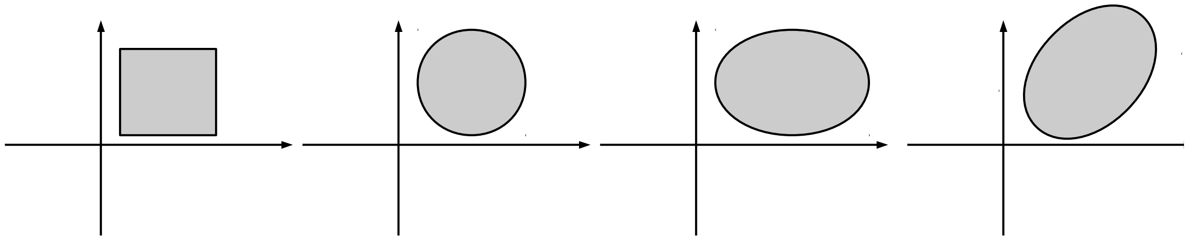


Figure 1: Joint PDF of X and Y is uniform in shaded areas.

i. Determine $E[Y]$ and $E[XY]$.
ii. Find $cov(X,Y)$. Are X and Y uncorrelated ?
iii. Are X and Y independent ? Explain.

c) Answer the following questions.

i. Show that $E[g(X)+Y|X=a] = g(a)+E[Y|X=a]$. Justify your steps.
ii. Show that $E[g(X)Y|X] = g(X)E[Y|X]$. Justify your steps.
iii. Evaluate $E[X^2+Y]$ if $E[Y|X]=X^3$ and X is uniformly distributed in $[0,2]$.

4. We consider the sum

$$Y = \sum_{i=1}^{N} X_i,$$

where N is a Poisson-distributed random variable with parameter a, and X_1, X_2,...,X_N are independent identically distributed (i.i.d.) random variables.

1. Find the moment generating function (MGF) of N, i.e., $M_N(s) = E\left\{e^{sN}\right\}$.
2. If X_i's are exponential r.v. with mean $\frac{1}{\lambda}$, find the MGF of Y, i.e., $M_Y(s) = E\left\{e^{sY}\right\}$.
3. If X_i's are Bernoulli r.v. with mean mean p, find $M_Y(s)$. Then, find the PDF of Y.

5. Let the random variable Z_n be

$$Z_n = \tfrac{1}{\sqrt{n}} \textstyle\sum_{\imath=1}^{n} \frac{X_\imath - \mu}{\sigma} \qquad \forall n \in \mathbb{Z}_+,$$

where μ is a real number, σ is a positive real number, and $X_1, X_2, \ldots$ are independent identically distributed random variables satisfying $\mathbf{E}[X_\imath] = \mu$ and $var(X_\imath) = \sigma^2$. The Central limit theorem (CLT) asserts that

$$\lim_{n\to\infty} F_{Z_n}(z) = \Phi(z) \qquad \forall z \in \mathbb{R}, \tag{1}$$

where $\Phi(z)$ is the cumulative distribution function (CDF) of the standard normal random variable, i.e., $\Phi(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$. The convergence in CLT is uniform in the sense that

$$\lim_{n\to\infty} \sup_z |F_{Z_n}(z) - \Phi(z)| = 0. \tag{2}$$

Furthermore, if X_1 has finite absolute third moment, i.e. $\mathbf{E}\left[|X_1|^3\right] = \rho < \infty$, then

$$\sup_z |F_{Z_n}(z) - \Phi(z)| \leq 0.48 \tfrac{\rho}{\sigma^3 \sqrt{n}} \qquad \forall n \in \mathbb{Z}_+, \tag{3}$$

by the Berry–Esseen theorem. It is worth noting that (3) is valid for any positive integer n and for any distribution of X_1 satisfying $\mathbf{E}[|X_1|] = \rho$ and $var(X_1) = \sigma^2$. The bound in (3) is not an asymptotic claim such as the ones in (1) and (2).

In the following, we observe the convergence described in (1), its uniformity described in (2), and bound on the rate of convergence described in (3) for Poisson random variables. Let $X_\imath$ be a unit mean Poisson random variable for all $\imath \in \mathbb{Z}_+$ Let us denote the set $\left\{\frac{k}{20} : k \in \{-100, 99, \ldots, 99, 100\}\right\}$ by A.

(a) Prove that $\sqrt{n}Z_n + n$ is a mean n Poisson random variable.

(b) Calculate and plot using MATLAB $F_{Z_1}(z)$ for all $z \in A$. *Remark:* Via "poisscdf_ee230.m" you may call the function "poisscdf_ee230(x,λ)" for a vector x and scalar λ, which calculates the CDF of a mean λ Poisson random variable at x.

(c) Calculate and plot on the same figure $F_{Z_{10}}(z)$ for all $z \in A$.

(d) Calculate and plot on the same figure $F_{Z_{100}}(z)$ for all $z \in A$.

(e) Calculate and plot on the same figure $F_{Z_{1000}}(z)$ for all $z \in A$.

(f) Calculate and plot on the same figure $\Phi(z)$ for all $z \in A$. *Hint:* Via "normcdf_ee230.m" you may call the function "normcdf_ee230(x)" for a vector x, which calculates the CDF of the standard normal random variable at x.

(g) For all $n \in \{1, 5, 10, 25, 50, 75, 100, 250, 500, 750, 1000\}$ calculate

$$\Delta(n) = \max_{z \in A} |F_{Z_n}(z) - \Phi(z)|,$$

plot $\Delta(n)$ vs $\ln(n)$, and observe that $\Delta(n) \downarrow 0$ with the increasing n.

(h) Plot $(\Delta(n)\sqrt{n})$ vs $\ln(n)$ and observe that one can find a $K \in \mathbb{R}_+$ such that $\Delta(n) \le \frac{K}{\sqrt{n}}$ at least for the values of n considered.