CSC411 Homework 1

Deadline: Wednesday, Sept. 26, at 11:59pm

Name: Baijiang, Yuan

Student Number: 1002175335

<u>Question 1(a)</u>

$\because x \sim Uniform(0,1)\ and\ y \sim Uniforma(0,1)$

$\therefore E(x) = E(y) = \int_b^a xf(x)dx = \int_0^1 xp(x)dx = \int_0^1 \frac{x}{1-0}dx = \frac{1}{2}(1-0) = \frac{1}{2}$

$\therefore E(x^2) = E(y^2) = \int_0^1 \frac{x^2}{1-0}dx = \frac{1}{3}(1-0) = \frac{1}{3}$

$\therefore \boldsymbol{E(Z)} = E((x-y)^2) = E(x^2 - 2xy + y^2) = E(x^2) - 2E(x)E(y) + E(y^2) = 2\times \frac{1}{3} - 2\times\frac{1}{4} = \frac{1}{6}$

$\because Var(Z) = E(z^2) - E(z)^2$

$\because E(z^2) = E(x^2 - 2xy + y^2)^2 = E(x^4 - 2x^3y + x^2y^2 - 2x^3y + 4x^2y^2 - 2xy^3 + x^2y^2 - 2xy^3 + y^4)$

$\because E(x^4) = E(y^4) = \int_0^1 x^4dx = \frac{1}{5}(1-0) = \frac{1}{5}$

$\because E(x^3) = E(y^3) = \int_0^1 x^3dx = \frac{1}{4}(1-0) = \frac{1}{4}$

$\because E(x^2) = E(y^2) = \frac{1}{3}\ (Caculated\ above)$

$\therefore E(z^2) = \frac{1}{5} - 2\cdot\frac{1}{4}\cdot\frac{1}{2} + \frac{1}{3}\cdot\frac{1}{3} + 2\cdot\frac{1}{4}\cdot\frac{1}{2} + 4\cdot\frac{1}{3}\cdot\frac{1}{3} - 2\cdot\frac{1}{2}\cdot\frac{1}{4} + \frac{1}{3}\cdot\frac{1}{3} - 2\cdot\frac{1}{2}\cdot\frac{1}{4} + \frac{1}{5} = \frac{1}{15}$

$\therefore \boldsymbol{Var(Z)} = E(z^2) - E(z)^2 = \frac{1}{15} - \left(\frac{1}{6}\right)^2 = \frac{7}{180}$

$\therefore The\ expecation\ and\ variance\ of\ Z = (x-y)^2\ is\ \frac{1}{6}\ and\ \frac{7}{180}\ respectively.$

<u>Question 1(b)</u>

$\mathbb{E}(\mathrm{R}) = \mathrm{E}[Z_1 + Z_2 + \cdots + Z_d] = \mathrm{E}(Z_1) + \cdots \mathrm{E}(Z_d) = \sum_{i=0}^{d} \frac{1}{6} = \frac{d}{6}$

$Var(R) = Var[Z_1 + Z_2 + \cdots + Z_d] = \sum_{i=0}^{d} Var(z_I) + \sum_{i\neq j}^{d} Cov(Z_i + Z_j)$

$\because x, y\ are\ independent. And,\ Z_i = (x_i + y_i)^2\ and\ Z_j = (x_j + y_j)^2$

$\therefore Z_i\ and\ Z_j\ are\ mutually\ independent. Therefore\ Cov(Z_i + Z_j) = 0$

$\therefore \boldsymbol{Var(R)} = \sum_{i=0}^{d} Var(z_I) + \sum_{i\neq j}^{d} Cov(Z_i + Z_j) = \sum_{i=0}^{d} \frac{7}{180} + 0 = \frac{7d}{180}$

Question 2(b)

I pick max_depth as 1,3,5,7,9

outputs are :

criterion='gini', max_depth =1,correct = 310,total=490,accuracy =0.6326530612244898

criterion='gini', max_depth =3,correct = 330,total=490,accuracy =0.673469387755102

criterion='gini', max_depth =5,correct = 327,total=490,accuracy =0.6673469387755102

criterion='gini', max_depth =7,correct = 337,total=490,accuracy =0.6877551020408164

criterion='gini', max_depth =9,correct = 348,total=490,accuracy =0.710204081632653

criterion='entropy', max_depth =1,correct = 310,total=490,accuracy =0.6326530612244898

criterion='entropy', max_depth =3,correct = 330,total=490,accuracy =0.673469387755102

criterion='entropy', max_depth =5,correct = 327,total=490,accuracy =0.6673469387755102

criterion='entropy', max_depth =7,correct = 330,total=490,accuracy =0.673469387755102

criterion='entropy', max_depth =9,correct = 335,total=490,accuracy =0.6836734693877551

Question 2(c)