1. a). $X \sim U(0,1)$, $Y \sim U(0,1)$

$\Rightarrow E(X) = \frac{1}{2}(a+b) \Rightarrow E(Y) = 0.5$

$= \frac{1}{2}(0+1)$ $\quad Var(Y) = 1/12$

$= 0.5$

$Var(X) = \frac{1}{12}(b-a)^2$

$= \frac{1}{12}(1-0)^2$

$= 1/12$

$\Rightarrow Z \equiv (X-Y)^2$

$E(Z) = E[(X-Y)^2]$

$= E(X^2 - 2XY + Y^2)$

$= E(X^2) - 2E(X)E(Y) + E(Y^2)$

but $Var(X) = E(X^2) - E^2(X)$

$E(X^2) = Var(X) + E^2(X)$

$= Var(X) + E(X)E(X) - 2E(X)E(Y) + Var(Y) + E(Y)E(Y)$

$= \frac{1}{12} + \frac{1}{4} - \frac{2}{4} + \frac{1}{12} + \frac{1}{4}$

$= \boxed{\frac{1}{6}}$

$Var(Z) = E(Z^2) - E^2(Z)$

$= E((X-Y)^4) - 1/36$

$= E(X^4 - 4X^3Y + 6X^2Y^2 - 4XY^3 + Y^4) - 1/36$

$= E(X^4) - 4E(X^3)E(Y) + 6E(X^2)E(Y^2) - 4E(X)E(Y^3) + E(Y^4) - 1/36$

$= 2E(X^4) - 8E(X^3)E(X) + 6E^2(X^2) - 1/36$

$$\Rightarrow E(x^p) = \frac{1}{p+1} \sum_{i=0}^{p} a^i b^{p-i}$$

$$E(x^2) = \frac{1}{2+1} \sum_{0}^{2} a^i b^{2-i}$$

$$= \frac{1}{3}\left(b^2 + \cancel{ab + a^2}\right)$$

$$= \frac{1}{3}$$

$$E(x^3) = \frac{1}{3+1} \sum_{0}^{3} a^i b^{3-i}$$

$$= \frac{1}{4}\left(b^3 + \cancel{ab^2 + ba^2 + a^3}\right)$$

$$= 1/4$$

$$E(x^4) = \frac{1}{5}\left(b^4 + \cancel{b^3a + b^2a^2 + ba^3 + a^4}\right)$$

$$= 1/5$$

$$\Rightarrow Var(z) = 2(1/5) - 8(1/2)(1/4) + 6(1/3)^2 - 1/36$$

$$= \boxed{7/180}$$

2. b). $E(R)$

$= E(z_1 + \dots + z_d)$

$= E\left((x_1 - y_1)^2 + \dots + (x_d - y_d)^2\right)$

$= E(z_1) + E(z_2) + E(z_3) + \dots + E(z_d)$

$= \frac{1}{6} + \frac{1}{6} + \frac{1}{6} + \dots + \frac{1}{6}$

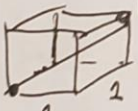
$= \boxed{\frac{d}{6}}$

$Var(R) = Var(z_1 + \dots + z_d)$

$= Var(z_1) + \dots + Var(z_d)$

$= \boxed{\frac{7d}{180}}$

2. c).



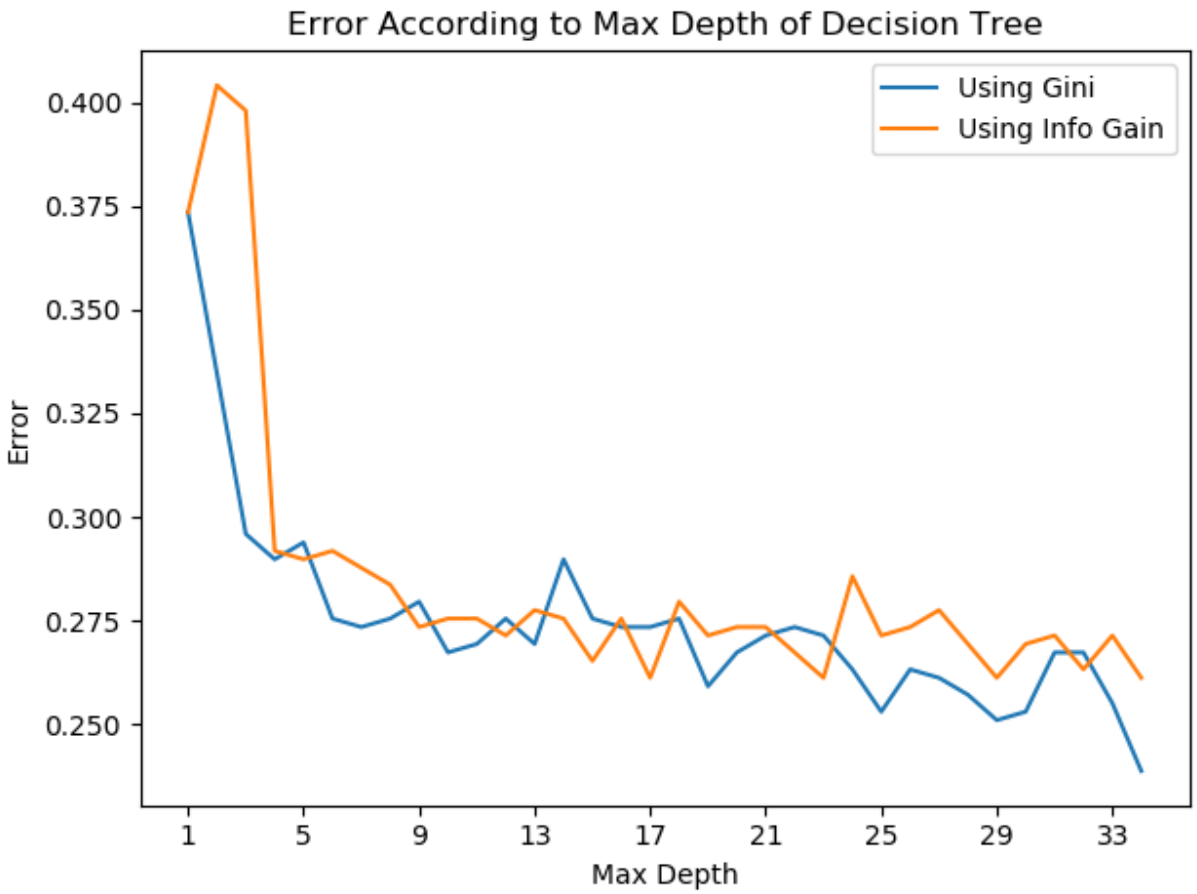
$\Rightarrow$ dist. $= \underbrace{1^2 + 1^2 + \dots + 1^2}_{d \text{ times.}}$

$= d$

$\sigma = \sqrt{\frac{7}{180}}\sqrt{d}$

-- in greater dimensions, $E(x) >> \sigma$ meaning 2 points have greater distances in between eachother and since there is a much less deviation in distances, they are approx. the same.

2. b)



Here are the accuracies for a max depth of 1 – 33 for each split criterion. I chose to use a Gini hyperparameter and max depth of 6 for part c).

Printout:

For depth 0, gini error = 0.34285714285714286, ig error = 0.34285714285714286

For depth 1, gini error = 0.33061224489795915, ig error = 0.34285714285714286

For depth 2, gini error = 0.30204081632653057, ig error = 0.30204081632653057

For depth 3, gini error = 0.30204081632653057, ig error = 0.30204081632653057

For depth 4, gini error = 0.310204081632653, ig error = 0.30204081632653057

For depth 5, gini error = 0.29387755102040813, ig error = 0.310204081632653

For depth 6, gini error = 0.28979591836734697, ig error = 0.28775510204081634

For depth 7, gini error = 0.2918367346938775, ig error = 0.28367346938775506

For depth 8, gini error = 0.28163265306122454, ig error = 0.28367346938775506

For depth 9, gini error = 0.27755102040816326, ig error = 0.28775510204081634

For depth 10, gini error = 0.2755102040816326, ig error = 0.28163265306122454

For depth 11, gini error = 0.27142857142857146, ig error = 0.2795918367346939

For depth 12, gini error = 0.273469387755102, ig error = 0.2857142857142857

For depth 13, gini error = 0.27755102040816326, ig error = 0.28163265306122454

For depth 14, gini error = 0.2795918367346939, ig error = 0.28367346938775506

For depth 15, gini error = 0.273469387755102, ig error = 0.28163265306122454

For depth 16, gini error = 0.2755102040816326, ig error = 0.2693877551020408

For depth 17, gini error = 0.2693877551020408, ig error = 0.2979591836734694

For depth 18, gini error = 0.27142857142857146, ig error = 0.27755102040816326

For depth 19, gini error = 0.2693877551020408, ig error = 0.2755102040816326

For depth 20, gini error = 0.2673469387755102, ig error = 0.27755102040816326

For depth 21, gini error = 0.2693877551020408, ig error = 0.2795918367346939

For depth 22, gini error = 0.2551020408163265, ig error = 0.273469387755102

For depth 23, gini error = 0.26530612244897955, ig error = 0.27755102040816326

For depth 24, gini error = 0.263265306122449, ig error = 0.2673469387755102

For depth 25, gini error = 0.273469387755102, ig error = 0.273469387755102

For depth 26, gini error = 0.26530612244897955, ig error = 0.27755102040816326

For depth 27, gini error = 0.263265306122449, ig error = 0.263265306122449

For depth 28, gini error = 0.2755102040816326, ig error = 0.2693877551020408

For depth 29, gini error = 0.2673469387755102, ig error = 0.273469387755102

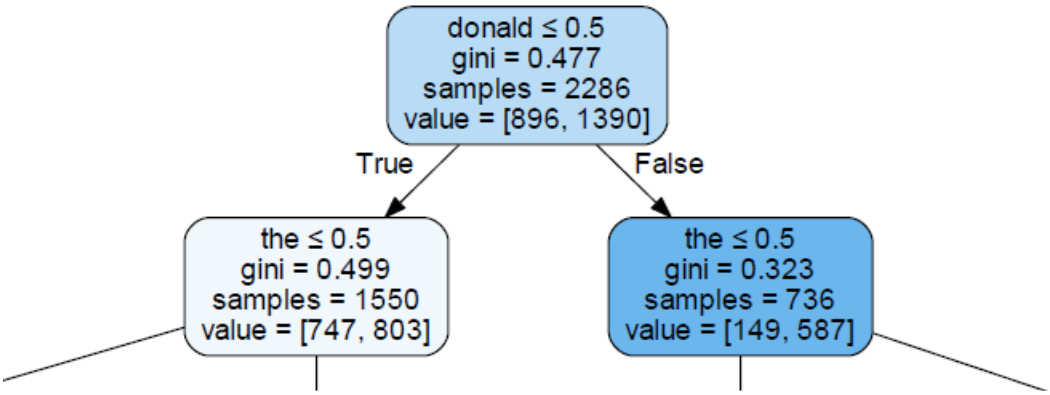
For depth 30, gini error = 0.2673469387755102, ig error = 0.2673469387755102

For depth 31, gini error = 0.2673469387755102, ig error = 0.263265306122449

For depth 32, gini error = 0.2673469387755102, ig error = 0.27142857142857146

For depth 33, gini error = 0.2571428571428571, ig error = 0.2571428571428571

2. c)



2. d)

Printout:

I(Y, donald) = 0.7303406536966675

I(Y, trumps) = 0.9609529245895325

I(Y, the) = 0.8317028852240703

I(Y, hillary) = 0.9459847151200842

I(Y, voting) = 0.966103516969252