**Question 1**

Test hypothesis h has errors r = 240 on sample S of n =800.

The standard deviation for

Example in class had and

Thus, standard deviation for

As no other information is given, most probable value of is this also assumes standard unbiased estimator for and data sample S is independent of discrete-valued hypothesis h.

**Question 2**

Sample S contains n = 100.

Incorrect classification, r = 100 – correct classification.

Error, r = 21.

95% confidence Implies,

Standard deviation

**Upper bound = 0.21+0.0798 = 0.130**

**Lower bound = 0.21 – 0.0798 = 0.2898**

**Question 3**

Range of interval, L = 0.3 and U = 0.6

Therefore, midpoint, M =

And (1 - p) = 1 – 0.45 = 0.55

Lower bound, upper bound,

Confidence interval width is

95% of two level confidence, we use Z0.975 = 1.96

Minimum number of examples

Thus minimum number of examples needed to collect should be **greater than or equal to 380**.

**Question 4**

90% 2-sided true error:

95% 1-sided true error:

…consider only Upper bound

0.0392

= 0.197

80% 1-sided true error:

…consider only Upper bound

**= 0.1831**

**Question 5**

Dimension for linear separator in 2-dimensional.

For Upper bound,

**This bound does not seem to be realistic.**

**Because a hyperplane in 2 dimension is a line which has to be defined by 2 points. So if 90% confidence with at most 5% error will break the line. So this might be the reason to sound unrealistic.**

**Question 6**

1. For rectangle,

Consider point in and and

Consider that 3 points are positive and a point is negative in the following case. Here it is not possible to shatter the points in.

Hence, the

1. For circle,

Points in circle are positive and outside is negative.

for circle in is at least 3, as 3 points make up non-degenerated triangle.

It is possible to shatter 1,2 or 3 positive points so,

**Question 7**

(a)

Region bounded by point (0,0) and (n,n) in the interval (0,100) implies that n = 101.

with probability 95% = 100 – 95 = 0.05

Size of hypothesis m is calculated as:

(b)

Region bounded by point (0,100) a read values,

**Question 8**

(a)

The tree has depth 2 and 4 leaves in total.

So, syntactically distinct trees are

( )

/ \

( ) ( )

/ \ / \

( ) ( ) ( ) ( )

(b)

Upper bound for number of examples, m with

**Question 9**

(a)

should be calculated to find the error.

…given

As , standard deviation = 0 and true error is also 0.

This implies that it is difficult to calculate the true error with 95% probability as

Hence it is difficult to calculate the true error since it is difficult to find the interval.

(b)

Upper bound = 0.3898 & Lower bound = 0.2101

**The interval in which this true error will fall is 0.2101 to 0.3898**