**Learning Theory – Exercise 1**

Question 1:

Section a:

First, observe that:

Since are a non-increasing rearrangement of .  
Given that, one may remember that are i.i.d. This leads to also being i.i.d, because the absolute value function on the above random variables series does not change its independency property and it is most certainly preserves the identical distribution of the random variables since the same mapping (or function) is applied on all of them.

Now, (**1**) can be written as:

Using **Chebyshev's inequality** as learned in class, we can come up with an upper bound on each of the probabilities inside the above product:

Note: Since the norm considers the **expectation of the absolute random variable** it is being applied on, the last equation holds. The reason is that is a copy of random variable and therefore being distributed samely.Also, one may notice that the following holds:

Since .

Using (**4**):

From (**2**) and (**5**):

Section b:

In this section we are dealing with a sub-case of section a, in which are independent and distributed as standard Gaussians.  
In order to show the wanted bound, one may assist the following definition of the norm:

Now, we can assist section (a) where we have come up with an upper bound on random variables which satisfy the conditions that does for all k:

Where is distributed as standard Gaussian random variable.

Placing (2) into (1):

Applying the constraint for convergence:

We want to form an inequality regarding the first moment of. Therefore, a reasonable choice of p would be.

As said, since where g is a standard Gaussian r.v.:

Let us show that there is a positive constant for which:

For

Question 3:

Section a:

Let us express the requested norm:

The above expectation is:

And our goal is to find some for which the smallest value of for which the above expectation is smaller or equal than 2 is .  
Since the expression inside the exponent involves cross multiplications of , one cannot integrate separately over each relative components.  
Thus, we will simply pick a vector and show the result this choice leads to.

Observe the vector . Meaning, a peaky vector in the space. Following this choice, we get:

Using the facts that:

are independent random variables.

takes only non-negative values in the above integration.

Now, for every absolute constant , the integral:

Does not converge.

On the other hand, for the above satisfies:

Clarification:  
We say that in the sense that, given the following problem:

Which we developed into:

Satisfies:

Thus, a direction in which is

Section b:

Now, we wish to find an absolute constant c that satisfies the following:

Given the definition of the norm:

Find c, such that:

For **all possible** .

Developing the norm expression:

Where the last equation holds due to the independency of and the triangle inequality. Now, Due to the integration borders, one can express the above as:

For the above expression not to converge, one must apply the restrictions:

The result of this integration is:

Now, remembering the definition of norm, the following inequality must hold:

So, now our goal is to find the absolute constant that is the maximum value of all the infimum values achieved by varying over all different sets of , and generating the corresponding norm as defined.  
(Under the constraint of course)  
Notice that since are distributed with a standard exponential distribution -.

Conclusively, to formulate the optimization problem, we look to find:

For all different values of .

Let's consider the case where .  
In that case:

The above final inequality leads us to the wanted absolute value in two ways.

First, remembering our optimization problem above, one's goal is to find the maximum value that the infimum of can have. Well, for a given value of it is easy to see that On the other hand, finding the maximum value of the above infimum for all possible value of is a problem of a different flavor.  
Notice that as long as the assumption that a uniform vector gives the wanted constant (we will address that in the next paragraph), one would like to put an upper estimate on at this point.  
Let be the vector which achieves the maximum value of norm out of all possible uniform vectors. Moreover, let be the norm results by choosing the direction .   
Then, choosing and using the achieved inequality gives the desired :

Second, it approves of the fact that is achieved by choosing a uniform vector in which the value of each element is the maximum possible. Let us assume that the chosen to achieve the absolute constant is a vector in which all the elements (or at least one of them) are smaller than .  
It can be verified by simple mathematical steps that if the exact same stages we have made thus far in order to come up with the are applied with , the resulted will hold . Meaning, the constant that answers the demands of our problem is achieved as seen, using the uniform .

Question 5:

Let us look at the definition of norm given the internal product argument:

We would like to show that the set:

Holds:

Where c is an absolute constant and **1** is the indicator function.