**Learning Theory – Exercise 1**

Question 1:

Section a:

First, note that:

Second, 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.  
Now, let us remember that the above analysis was considering only a very specific selection of k measurements. On the other hand, there different possibilities of that sort. Using that information:

Finally, all that has left is using the upper bound on the binomial coefficients:

Leading to

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.  
Before our calculations, let us define the value of above as .

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   
In order to do so, let us rephrase the problem in question. We want to find the maximum value of the expression:

Let us use the connection between n and k and set .  
Thus, the expression becomes:

One may notice that this is a monotonically non-decreasing function of the variable n. Meaning, analyzing the limit will yield the requested maximum value.   
If one re-arranges the expression to leave out all values which are independent in n:

Using standard techniques from calculus courses we can yield that:

Finally, we can conclude that:

Given.

**TBD**

Note, this expression holds for all . For we can analyze the original inequality:

One can observe that:

Section c:

Question 2:

Section a:

The norm of vector is defined as .  
Likely, the expectation on that norm is .  
Our objective is to find the best upper and lower bounds on the above given that are distributed as independent, standard Gaussians.

Let us start with the lower bound using Jensen's inequality:

Since expectation is a linear operator, one can change the order of the sum and the expectation:

One naïve way to bound the above from below is by noticing that:

And so:

Now, let us look at the other side of the inequality:

Meaning that now our objective is finding how is distributed.  
Let us define .  
In that case:

Using the fact that our Gaussian vector is standard and its covariance matrix is diagonal, we may infer that:

Where we used folded normal distribution with standard parameters.  
Now:

And this result is the wanted distribution. Now our goal is to find the expectation of the above, which is given by:

And so the upper bound is given by:

And to sum up:

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 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 -.

Now, we will use the fact that , meaning that.

In that case:

Now, it is obvious that

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.

Let us start by analyzing the following:

Solving this inequality with respect to yields:

Meaning, given the length of our vector, n:

To conclude, we showed that for given vectors as defined, the absolute constant c that holds:

Satisfies that:

For all possibilities of the vector , and so specifically, for more than possibilities of it.