Re: Question relating to order statistics of normal variables

Source: http://sci.tech−archive.net/Archive/sci.stat.edu/2006−09/msg00004.html

- *From*: J W <julianw@xxxxxxxxxxxxxxxxxxxxx>
- *Date*: Sat, 02 Sep 2006 23:50:15 −0700

shrishri@xxxxxxxxx wrote:

> Hi all,
>
> I am looking for a solution / direction to solve the following problem.
>
> Any pointers will be greatly appreciated!
>
> The problem I am seeking a solution for is as follows.
> Let X_1, X_2,... X_N be N i.i.d Gaussian random variables with unit
> variance and zero mean.
> Suppose we arrange the realizations in increasing values of their
> absolute values
> (similar to order statistics, except that we are using the absolute
> values to sort them), and let the corresponding sequence of
> random variables be Y_1, Y_2,...,Y_N.
> I am looking for the expected value of sum of squares of the first K
> terms of the sequence Y_i, in the asymptotic regime where
> N tends to infinity, and K/N is a constant in the open interval (0,1).
>
> If someone has a pointer to the literature, that would be great!
> Thanks −Shri.

Don't know the literature in this area, and I can't guarantee I'm getting you any further than you've already gotten (or that it's even right − it's late, and I'm studying for qualifiers), but here's something:

First, note that the Y_i's are half−normal. Then you can get the joint distribution of the order statistics for a half normal using the standard n! prod (f(y_i)) 1 [y_1 < ... < y_n] formula. So now you're left with finding E(Y_1^2 + ... + Y_k^2), where the expectation is taken with respect to the joint distribution of the half−normal order statistics. This, of course, equals E(Y_1^2) + ... + E(Y_k^2). All the terms are of similar form − take E(Y_1^2), for example:

E(Y_1^2) ~= int_0^{y_1} x^2 exp(−x^2 / 2 s^2) * prod_{i=2 to k} int_{y_{i−1}}^k exp(−y_i^2 / 2 s^2)

(I'm leaving out some constants here)

The first part is just the second moment of a half−normal (you can find that), and the second part is a product

of F(y_i) − F(y_{i−1}) for i = 2 to k, where F is the half−normal CDF (has to do with the "erf", according to Mathworld).

That's as far as I got; maybe if you put things together, you can get somewhere. In any case, hope it helps a bit.

−J
.