Anirban's Angle: Top Inequalities for a PhD student

Contributing Editor Anirban DasGupta writes:

It is the mark of an instructed mind, said Aristotle, not to seek exactness when only an approximation of the truth is possible. Delicate and classy, still, the nature of mathematics is such that quantities of intrinsic importance often cannot be evaluated in a simple or explicit form. So one opts for the next best thing. Bound it from above or below by something simpler and explicit.

Inequalities form an integral part of the theory and practice of mathematical sciences. One we see in high school is that the irrational number $\pi < \frac{22}{7}$, as $\frac{22}{7} - \pi = \int_0^1 \frac{x^4(1-x)^4}{1+x^2}dx > 0$. For statisticians, that $P(\cup_{i=1}^n A_i) \leq \sum_{i=1}^n P(A_i) - \frac{2}{n}\sum_{i<j} P(A_i \cap A_j)$. There are countless inequalities; some are beautiful, some highly useful, some both. Which ones should a PhD student in mathematical statistics know?

To get a finger on my colleagues' pulse, I took a small poll. I asked Saugata Basu, Rabi Bhattacharya, Burgess Davis, Peter Hall, Iain Johnstone, B.V. Rao, Yosi Rinott, Philip Stark, Sara van de Geer, and Jon Wellner. Of course, the choices differed. As an experiment in innocuous merriment, I chose my favorites. My collection is embarrassingly biased by at least three factors: inequalities that I at least know, those I have personally seen being applied, and liked—either the application or the inequality itself.

My one-page limit keeps me from stating all the inequalities, and so I only mention them by name or descriptively. Perhaps it would be useful to have them precisely stated, proved, each illustrated with one good application, and made publicly available in some platform.

Here then, cerebrating, is a list of inequalities I would wish to know, if I were a graduate student working on statistical theory today. They are generally grouped by topics; analysis, matrices, probability, moments, limit theorems, statistics.

1 Cauchy–Schwarz
2 Jensen
3 Hölder and triangular
4 Fatou
5 Bessel
6 Hausdorff–Young
7 Basic Sobolev inequality in three dimensions only
8 Frobenius
9 Sylvestre
10 Determinant bounds, e.g., Hadamard
11 Kantorovich
12 Courant–Fischer
13 Boole's inequality, from both directions
14 Chebyshev and Markov
15 Bernstein
16 Hoeffding in the Rademacher case, 1963
17 Bounds on Mills ratio from both directions
18 Upper tail of Binomial and Poisson
19 Slepian's lemma, 1962
20 Anderson's inequality on probabilities of symmetric convex sets, 1955
21 Rosenthal, 1970
22 Kolmogorov's basic maximal inequality
23 Basic Berry-Esseen in one dimension
24 Le Cam's bound on Poisson approximations (Le Cam, 1960)
25 DKW with a mention of Massart's constant (Massart, 1990)
26 Bounds on expectation of normal maximum from both directions
27 Comparison lemma on multinormal CDFs (Leadbetter, Lindgren, and Rootzén, 1983)
28 Talagrand (as in 1995, Springer)
29 Inequality between Hellinger and Kullback–Leibler distance
30 Cramér-Rao
31 Rao–Blackwell (which is an inequality)
32 Wald's SPRT inequalities.

Truly going back to my student days, I recall how useful matrix inequalities were in that period, when linear inference was such an elephant in the room. Inequalities on CLTs and metrics played pivotal roles in the sixties, and then again, as the bootstrap and later, MCMC, emerged. Concentration inequalities came to the forefront with the advent of empirical process theory, and then as high dimensional problems became important. It seems as though the potential of analytic inequalities in solving statistical and probabilistic problems hasn't yet been efficiently tapped. The recent book by Peter Bühlmann and Sara van de Geer (2011) has many modern powerful inequalities. There are of course new editions of the classics, e.g., Hardy, Littlewood and Pólya (1988), Marshall, Olkin and Arnold (2011).

Quite possibly, on another day I would include some other phenomenal inequalities, and drop some that I chose today. Can anyone vouch that Efron–Stein (1981), Gauss (for unimodal distributions), FKG (Fortuin, Kasteleyn, Ginibre, 1971), Chernoff's variance inequality (1981), or a basic prophet or log-Sobolev inequality, or even a basic Poincaré, need not be in the essential list? Defining what is the most useful or the most beautiful is about the most hopeless task one can have. Beauty and use are such indubitably personal choices. We have, in front of us, an ocean of remarkable inequalities. *You can't cross the sea*, said Nobel Laureate Poet Tagore, *merely by standing and staring at the water*. I figure I need to jump!