



**RUTGERS**  
UNIVERSITY

Department of Statistics & Biostatistics  
Hill Center, Room 501  
School of Arts & Sciences  
Rutgers, The State University of New Jersey  
110 Frelinghuysen Road.  
Piscataway, New Jersey 08854-8019

www.stat.rutgers.edu  
office@stat.rutgers.edu  
848-445-2690  
Fax: 732-445-3428

**RUTGERS UNIVERSITY**  
**DEPARTMENT OF STATISTICS AND BIOSTATISTICS**  
[www.stat.rutgers.edu](http://www.stat.rutgers.edu)

**Seminar**

Speaker: **Professor Hannes Leeb**  
**University of Vienna**

Title: **On the conditional distributions of low-dimensional projections from high-dimensional data**

Time: **1:00 – 2:00pm, Thursday, April 11, 2013**

Place: **552 Hill Center**

**Abstract**

We study the conditional distribution of low-dimensional projections from high-dimensional data, where the conditioning is on other low-dimensional projections. To fix ideas, consider a random  $d$ -vector  $Z$  that has a Lebesgue density and that is standardized so that  $\mathbb{E} Z = 0$  and  $\mathbb{E} Z Z' = I_d$ . Moreover, consider two projections defined by unit-vectors  $\alpha$  and  $\beta$ , namely a response  $y = \alpha' Z$  and an explanatory variable  $x = \beta' Z$ . It has long been known that the conditional mean of  $y$  given  $x$  is approximately linear in  $x$ , under some regularity conditions; cf. Hall and Li (1993). However, a corresponding result for the conditional variance has not been available so far. We here show that the conditional variance of  $y$  given  $x$  is approximately constant in  $x$  (again, under some regularity conditions). These results hold uniformly in  $\alpha$  and for most  $\beta$ 's, provided only that the dimension of  $Z$  is large. In that sense, we see that most linear submodels of a high-dimensional overall model are approximately correct. Our findings provide new insights in a variety of modeling scenarios. We discuss several examples, including sliced inverse regression, sliced average variance estimation, generalized linear models under potential link violation, and sparse linear modeling.

**\*\* Refreshments will be served at @12:40pm in Room 502 Hill Center \*\***