

## Sélection de variables avec données manquantes

Avner Bar-Hen<sup>1</sup>

<sup>1</sup> CNAM, Paris, France

Abstract : Les données manquantes sont une question récurrente en statistique et particulièrement dans le cas de la grande dimension. Nous considérons le cas d'une réponse  $Y$  et nous recherchons quelles variables parmi un ensemble de  $X_i$  sont significativement liées à la réponse  $Y$ . Les données manquantes (parmi les  $X_i$ ) sont considérées dans un cadre MCAR. Les propriétés de la procédure sont étudiées dans un cadre théorique et sur simulations. Une application sur données réelles permet de voir l'intérêt de la méthode.