**Generalized Linear Model for Insurance Data**

**By Piet de Jong and Gillian Z. Heller**

**DATA: Car**

This data set is based on one-year vehicle insurance

policies taken out in 2004 or 2005. There are 67856 policies, of

which 4624 (6.8%) had at least one claim.

Variables:

veh\_value vehicle value, in $10,000s

exposure 0-1

clm occurrence of claim (0 = no, 1 = yes)

numclaims number of claims

claimcst0 claim amount (0 if no claim)

veh\_body vehicle body, coded as

BUS

CONVT = convertible

COUPE

HBACK = hatchback

HDTOP = hardtop

MCARA = motorized caravan

MIBUS = minibus

PANVN = panel van

RDSTR = roadster

SEDAN

STNWG = station wagon

TRUCK

UTE - utility

veh\_age age of vehicle: 1 (youngest), 2, 3, 4

gender gender of driver: M, F

area driver's area of residence: A, B, C, D, E, F

agecat driver's age category: 1 (youngest), 2, 3, 4, 5, 6