# Multicollinearity

**STAT 615 (Baron)**

```
DATA C;
INFILE forest;
INPUT AGE HD N MDBH;

PROC REG DATA=C;
MODEL MDBH = HD AGE N / COLLIN TOL VIF;

RUN;
```

The REG Procedure
Model: MODEL1
Dependent Variable: MDBH

Number of Observations Read 20
Number of Observations Used 20

Analysis of Variance

| Source | DF | Sum of Squares | Mean Square | F Value | Pr > F |
|---|---|---|---|---|---|
| Model | 3 | 8.60560 | 2.86853 | 25.79 | <.0001 |
| Error | 16 | 1.77990 | 0.11124 | | |
| Corrected Total | 19 | 10.38550 | | | |

| | | | |
|---|---|---|---|
| Root MSE | 0.33353 | R-Square | 0.8286 |
| Dependent Mean | 6.26500 | Adj R-Sq | 0.7965 |
| Coeff Var | 5.32374 | | |

Parameter Estimates

| Variable | DF | Parameter Estimate | Standard Error | t Value | Pr > \|t\| | Tolerance | Variance Inflation |
|---|---|---|---|---|---|---|---|
| Intercept | 1 | 5.07300 | 0.52784 | 9.61 | <.0001 | . | 0 |
| HD | 1 | 0.09968 | 0.01873 | 5.32 | <.0001 | 0.18638 | 5.36536 |
| AGE | 1 | -0.09741 | 0.05244 | -1.86 | 0.0817 | 0.19665 | 5.08522 |
| N | 1 | -0.00290 | 0.00062451 | -4.64 | 0.0003 | 0.80108 | 1.24831 |

$1 - R_j^2$ → $VIF_j = \frac{1}{1-R_j^2}$

Collinearity Diagnostics

| Number | Eigenvalue | Condition Index | Proportion of Variation: Intercept | HD | AGE | N |
|---|---|---|---|---|---|---|
| 1 | 3.93496 | 1.00000 | 0.00127 | 0.00044404 | 0.00049800 | 0.00191 |
| 2 | 0.04558 | 9.29148 | 0.01797 | 0.02137 | 0.05106 | 0.34816 |
| 3 | 0.01575 | 15.80776 | 0.87202 | 0.04876 | 0.00145 | 0.38388 |
| 4 | 0.00372 | 32.53216 | 0.10874 | 0.92942 | 0.94700 | 0.26604 |

Eigenvalues of $(X'X)$, ordered

$\sqrt{\frac{\lambda_{max}}{\lambda_j}}$

$p_{je} = \frac{A_{je}^2 / \lambda_e}{VIF_j}$

Condition number: $\kappa = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$

where
$A$ = matrix that orthogonalizes $X'X$
$A_{je}$ = principal components
Scaled $X'X$ is used: all diagonals = 1