Model Selection

Taken from slides by OpenIntro Statistics and Professor Papachristou.

Beauty in the classroom

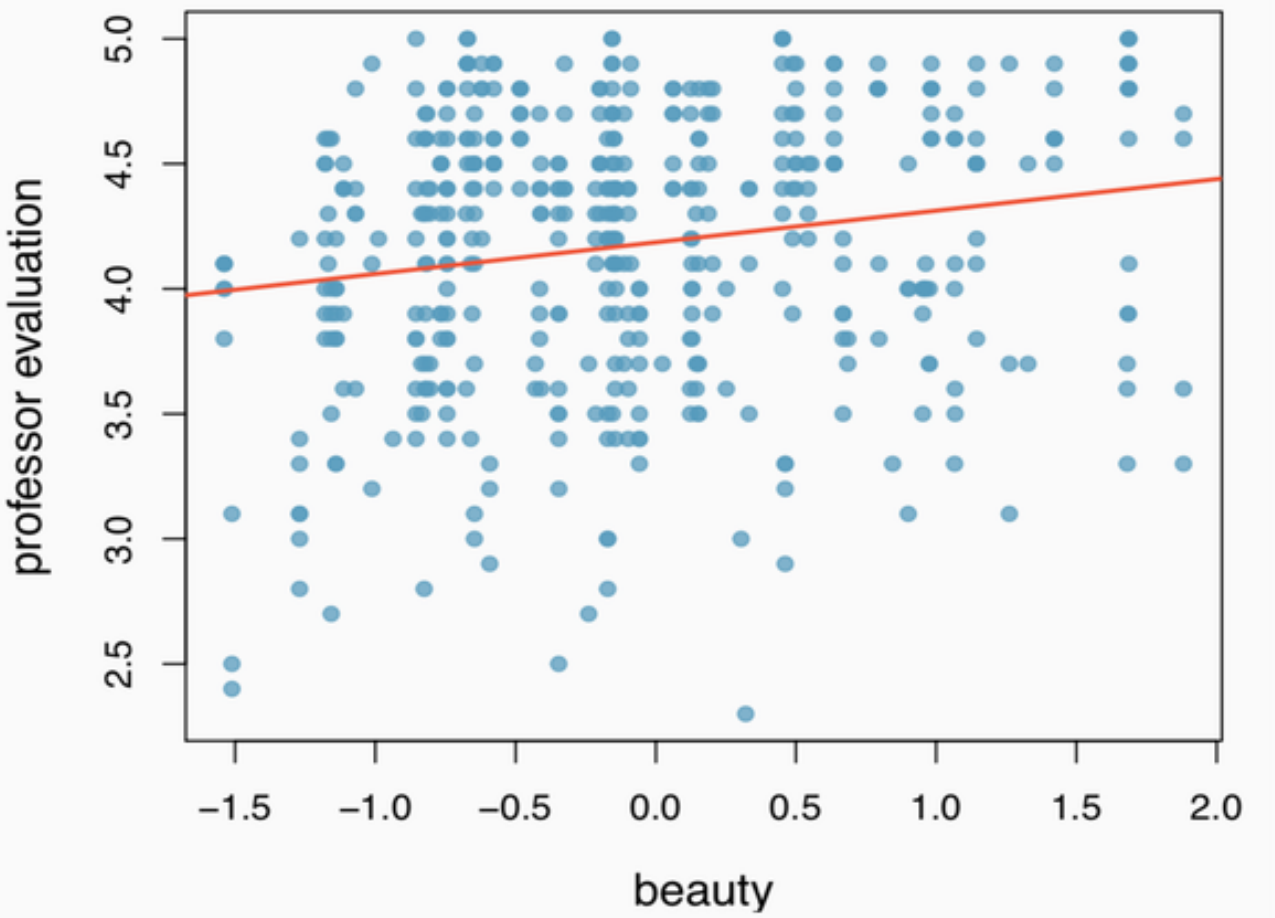
Data: Student evaluations of instructors' beauty and teaching quality for 463 courses at the University of Texas.

Evaluations conducted at the end of semester, and the beauty judgements were made later, by six students who had not attended the classes and were not aware of the course evaluations (2 upper level females, 2 upper level males, one lower level female, one lower level male).

Hamermesh & Parker. (2004) "Beauty in the classroom: instructors' pulchritude and putative pedagogical productivity" *Economics Education Review*.

Professor rating vs. beauty

Professor evaluation score (higher score means better) vs. beauty score (a score of 0 means average, negative score means below average, and a positive score above average):



Which of the below is correct based on the model output?

	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	4.19	0.03	167.24	0.00
beauty	0.13	0.03	4.00	0.00

$$R^2 = 0.0336$$

(a) Model predicts 3.36% of professor ratings correctly.

(b) Beauty is not a significant predictor of professor evaluation.

(c) Professors who score 1 point above average in their beauty score tend to also score 0.13 points higher in their evaluation.

(d) 3.36% of variability in beauty scores can be explained by professor evaluation.

(e) The correlation coefficient could be $\sqrt{0.0336} = 0.18$ or -0.18, we can't tell which is correct.

Which of the below is correct based on the model output?

	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	4.19	0.03	167.24	0.00
beauty	0.13	0.03	4.00	0.00

$$R^2 = 0.0336$$

(a) Model predicts 3.36% of professor ratings correctly.

(b) Beauty is not a significant predictor of professor evaluation.

(c) Professors who score 1 point above average in their beauty score tend to also score 0.13 points higher in their evaluation.

(d) 3.36% of variability in beauty scores can be explained by professor evaluation.

(e) The correlation coefficient could be $\sqrt{0.0336} = 0.18$ or -0.18, we can't tell which is correct.

Professor rating vs. beauty + gender

For a given beauty score, are male professors evaluated higher, lower, or about the same as female professors?

	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	4.09	0.04	107.85	0.00
beauty	0.14	0.03	4.44	0.00
gender.male	0.17	0.05	3.38	0.00

$$R^2_{adj} = 0.057$$

(a) higher

(b) lower

(c) about the same

Professor rating vs. beauty + gender

For a given beauty score, are male professors evaluated higher, lower, or about the same as female professors?

	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	4.09	0.04	107.85	0.00
beauty	0.14	0.03	4.44	0.00
gender.male	0.17	0.05	3.38	0.00

$$R^2_{adj} = 0.057$$

(a) higher $\rightarrow$ Beauty held constant, male professors are rated 0.17 points higher on average than female professors.

(b) lower

(c) about the same

Assessing significance: categorical variables

Tenure is a categorical variable with 3 levels: non-tenure track, tenure track, tenured. How to encode this into dummy variables with reference level being non-tenure track?

Original Data
Tenure Status
non-tenure track
tenure track
tenured
tenure track
non-tenure track
tenured

Assessing significance: categorical variables

Tenure is a categorical variable with 3 levels: non-tenure track, tenure track, tenured. How to encode this into dummy variables with reference level being non-tenure track?

Original Data	**Dummy Encoded (x_1, x_2)**	
Tenure Status	**tenure_track** (x_1)	**tenured** (x_2)
non-tenure track	0	0
tenure track	1	0
tenured	0	1
tenure track	1	0
non-tenure track	0	0
tenured	0	1

Assessing significance: categorical variables

Tenure is a categorical variable with 3 levels: non-tenure track, tenure track, tenured. How to encode this into dummy variables with reference level being non-tenure track?

Original Data	**Dummy Encoded** (x_1, x_2)	
Tenure Status	**tenure_track** (x_1)	**tenured** (x_2)
non-tenure track	0	0
tenure track	1	0
tenured	0	1
tenure track	1	0
non-tenure track	0	0
tenured	0	1

The reference **non-tenure track** is coded as $x_1 = 0$ and $x_2 = 0$.

Assessing significance: categorical variables

Tenure is a categorical variable with 3 levels: non-tenure track, tenure track, tenured. Based on the model output given, which of the below is false?

	Estimate	Std. Error	t value	$Pr(>\|t\|)$
tenure.tenure track	-0.1933	0.0847	-2.28	0.02
tenure.tenured	-0.1574	0.0656	-2.40	0.02

(a) Reference level is non-tenure track.

(b) All else being equal, tenure track professors are rated, on average, 0.19 points lower than non-tenure track professors.

(c) All else being equal, tenured professors are rated, on average, 0.16 points lower than non-tenure track professors.

(d) All else being equal, there is a significant difference between the average ratings of tenure track and tenured professors.

Assessing significance

Which predictors do not seem to meaningfully contribute to the model, i.e. may not be significant predictors of professor's rating score?

	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	4.6282	0.1720	26.90	0.00
beauty	0.1080	0.0329	3.28	0.00
gender.male	0.2040	0.0528	3.87	0.00
age	-0.0089	0.0032	-2.75	0.01
formal.yes	0.1511	0.0749	2.02	0.04
lower.yes	0.0582	0.0553	1.05	0.29
native.non english	-0.2158	0.1147	-1.88	0.06
minority.yes	-0.0707	0.0763	-0.93	0.35
students	-0.0004	0.0004	-1.03	0.30
tenure.tenure track	-0.1933	0.0847	-2.28	0.02
tenure.tenured	-0.1574	0.0656	-2.40	0.02

Assessing significance

Which predictors do not seem to meaningfully contribute to the model, i.e. may not be significant predictors of professor's rating score?

	Estimate	Std. Error	t value	Pr($>\|t\|$)
(Intercept)	4.6282	0.1720	26.90	0.00
beauty	0.1080	0.0329	3.28	0.00
gender.male	0.2040	0.0528	3.87	0.00
age	-0.0089	0.0032	-2.75	0.01
formal.yes	0.1511	0.0749	2.02	0.04
lower.yes	0.0582	0.0553	1.05	0.29
native.non english	-0.2158	0.1147	-1.88	0.06
minority.yes	-0.0707	0.0763	-0.93	0.35
students	-0.0004	0.0004	-1.03	0.30
tenure.tenure track	-0.1933	0.0847	-2.28	0.02
tenure.tenured	-0.1574	0.0656	-2.40	0.02

Model selection strategies

Based on what we've learned so far, what are some ways you can think of that can be used to determine which variables to keep in the model and which to leave out?

Backward-elimination

1. Start with the full model.
2. Drop one variable at a time and record R^2_{adj} of each smaller model.
3. Pick the model with the highest increase in R^2_{adj}.
4. Repeat until none of the models yield an increase in R^2_{adj}.

Backward-elimination

Full	beauty + gender + age + formal + lower + native + minority + students + tenure	0.0839
Step 1	gender + age + formal + lower + native + minority + students + tenure	0.0642
	beauty + age + formal + lower + native + minority + students + tenure	0.0557
	beauty + gender + formal + lower + native + minority + students + tenure	0.0706
	beauty + gender + age + lower + native + minority + students + tenure	0.0777
	beauty + gender + age + formal + native + minority + students + tenure	0.0837
	beauty + gender + age + formal + lower + minority + students + tenure	0.0788
	beauty + gender + age + formal + lower + native + students + tenure	0.0842
	beauty + gender + age + formal + lower + native + minority + tenure	0.0838
	beauty + gender + age + formal + lower + native + minority + students	0.0733

Backward-elimination

Full	beauty + gender + age + formal + lower + native + minority + students + tenure	0.0839
Step 1	gender + age + formal + lower + native + minority + students + tenure	0.0642
	beauty + age + formal + lower + native + minority + students + tenure	0.0557
	beauty + gender + formal + lower + native + minority + students + tenure	0.0706
	beauty + gender + age + lower + native + minority + students + tenure	0.0777
	beauty + gender + age + formal + native + minority + students + tenure	0.0837
	beauty + gender + age + formal + lower + minority + students + tenure	0.0788
	beauty + gender + age + formal + lower + native + students + tenure	0.0842
	beauty + gender + age + formal + lower + native + minority + tenure	0.0838
	beauty + gender + age + formal + lower + native + minority + students	0.0733
Step 2	gender + age + formal + lower + native + students + tenure	0.0647
	beauty + age + formal + lower + native + students + tenure	0.0543
	beauty + gender + formal + lower + native + students + tenure	0.0708
	beauty + gender + age + lower + native + students + tenure	0.0776
	beauty + gender + age + formal + native + students + tenure	0.0846
	beauty + gender + age + formal + lower + native + tenure	0.0844
	beauty + gender + age + formal + lower + native + students	0.0725

Backward-elimination

Full	beauty + gender + age + formal + lower + native + minority + students + tenure	0.0839
Step 1	gender + age + formal + lower + native + minority + students + tenure	0.0642
	beauty + age + formal + lower + native + minority + students + tenure	0.0557
	beauty + gender + formal + lower + native + minority + students + tenure	0.0706
	beauty + gender + age + lower + native + minority + students + tenure	0.0777
	beauty + gender + age + formal + native + minority + students + tenure	0.0837
	beauty + gender + age + formal + lower + minority + students + tenure	0.0788
	beauty + gender + age + formal + lower + native + students + tenure	0.0842
	beauty + gender + age + formal + lower + native + minority + tenure	0.0838
	beauty + gender + age + formal + lower + native + minority + students	0.0733
Step 2	gender + age + formal + lower + native + students + tenure	0.0647
	beauty + age + formal + lower + native + students + tenure	0.0543
	beauty + gender + formal + lower + native + students + tenure	0.0708
	beauty + gender + age + lower + native + students + tenure	0.0776
	beauty + gender + age + formal + native + students + tenure	0.0846
	beauty + gender + age + formal + lower + native + tenure	0.0844
	beauty + gender + age + formal + lower + native + students	0.0725
Step 3	gender + age + formal + native + students + tenure	0.0653
	beauty + age + formal + native + students + tenure	0.0534
	beauty + gender + formal + native + students + tenure	0.0707
	beauty + gender + age + native + students + tenure	0.0786
	beauty + gender + age + formal + students + tenure	0.0756
	beauty + gender + age + formal + native + tenure	0.0855
	beauty + gender + age + formal + native + students	0.0713

Backward-elimination

Full	beauty + gender + age + formal + lower + native + minority + students + tenure	0.0839
Step 1	gender + age + formal + lower + native + minority + students + tenure	0.0642
	beauty + age + formal + lower + native + minority + students + tenure	0.0557
	beauty + gender + formal + lower + native + minority + students + tenure	0.0706
	beauty + gender + age + lower + native + minority + students + tenure	0.0777
	beauty + gender + age + formal + native + minority + students + tenure	0.0837
	beauty + gender + age + formal + lower + minority + students + tenure	0.0788
	beauty + gender + age + formal + lower + native + students + tenure	0.0842
	beauty + gender + age + formal + lower + native + minority + tenure	0.0838
	beauty + gender + age + formal + lower + native + minority + students	0.0733
Step 2	gender + age + formal + lower + native + students + tenure	0.0647
	beauty + age + formal + lower + native + students + tenure	0.0543
	beauty + gender + formal + lower + native + students + tenure	0.0708
	beauty + gender + age + lower + native + students + tenure	0.0776
	beauty + gender + age + formal + native + students + tenure	0.0846
	beauty + gender + age + formal + lower + native + tenure	0.0844
	beauty + gender + age + formal + lower + native + students	0.0725
Step 3	gender + age + formal + native + students + tenure	0.0653
	beauty + age + formal + native + students + tenure	0.0534
	beauty + gender + formal + native + students + tenure	0.0707
	beauty + gender + age + native + students + tenure	0.0786
	beauty + gender + age + formal + students + tenure	0.0756
	beauty + gender + age + formal + native + tenure	0.0855
	beauty + gender + age + formal + native + students	0.0713
Step 4	gender + age + formal + native + tenure	0.0667
	beauty + age + formal + native + tenure	0.0553
	beauty + gender + formal + native + tenure	0.0723
	beauty + gender + age + native + tenure	0.0806
	beauty + gender + age + formal + tenure	0.0773
	beauty + gender + age + formal + native	0.0713

Backward-elimination

Step 3	gender + age + formal + native + students + tenure	0.0653
	beauty + age + formal + native + students + tenure	0.0534
	beauty + gender + formal + native + students + tenure	0.0707
	beauty + gender + age + native + students + tenure	0.0786
	beauty + gender + age + formal + students + tenure	0.0756
	beauty + gender + age + formal + native + tenure	0.0855
	beauty + gender + age + formal + native + students	0.0713
Step 4	gender + age + formal + native + tenure	0.0667
	beauty + age + formal + native + tenure	0.0553
	beauty + gender + formal + native + tenure	0.0723
	beauty + gender + age + native + tenure	0.0806
	beauty + gender + age + formal + tenure	0.0773
	beauty + gender + age + formal + native	0.0713

The R^2_{adj} values in Step 4 don't improve, so we are done at Step 3.

Backward-elimination

Step 3	gender + age + formal + native + students + tenure	0.0653
	beauty + age + formal + native + students + tenure	0.0534
	beauty + gender + formal + native + students + tenure	0.0707
	beauty + gender + age + native + students + tenure	0.0786
	beauty + gender + age + formal + students + tenure	0.0756
	beauty + gender + age + formal + native + tenure	0.0855
	beauty + gender + age + formal + native + students	0.0713
Step 4	gender + age + formal + native + tenure	0.0667
	beauty + age + formal + native + tenure	0.0553
	beauty + gender + formal + native + tenure	0.0723
	beauty + gender + age + native + tenure	0.0806
	beauty + gender + age + formal + tenure	0.0773
	beauty + gender + age + formal + native	0.0713

The R^2_{adj} values in Step 4 don't improve, so we are done at Step 3.

Forward-selection

1. Start with regressions of response vs. each explanatory variable.
2. Pick the model with the highest R^2_{adj}.
3. Add the remaining variables one at a time to the existing model, and once again pick the model with the highest R^2_{adj}.
4. Repeat until the addition of any of the remaining variables does not result in a higher R^2_{adj}.

Backward-Elimination vs. Forward-Selection

Backward elimination with p-value approach:

1. Start with full model.
2. Drop variable with highest p-value and refit smaller model.
3. Repeat until all variables left in model are significant.

Forward selection with p-value approach:

1. Start with regressions vs. each variable.
2. Pick variable with lowest significant p-value.
3. Add remaining variables one at a time, pick lowest significant p-value.
4. Repeat until no remaining variables have a significant p-value.

Adjusted R^2 vs. p-value approaches

The two approaches are similar, but they sometimes lead to different models, with the adjusted R^2 approach tending to include more predictors in the final model.

Adjusted R^2 vs. p-value approaches

The two approaches are similar, but they sometimes lead to different models, with the adjusted R^2 approach tending to include more predictors in the final model.

- When the sole goal is to improve prediction accuracy, use R^2. This is commonly the case in machine learning applications.

Adjusted R^2 vs. p-value approaches

The two approaches are similar, but they sometimes lead to different models, with the adjusted R^2 approach tending to include more predictors in the final model.

- When the sole goal is to improve prediction accuracy, use R^2. This is commonly the case in machine learning applications.
- When we care about understanding which variables are statistically significant predictors, or producing a simpler model, the p-value approach is preferred.

Adjusted R^2 vs. p-value approaches

The two approaches are similar, but they sometimes lead to different models, with the adjusted R^2 approach tending to include more predictors in the final model.

- When the sole goal is to improve prediction accuracy, use R^2. This is commonly the case in machine learning applications.
- When we care about understanding which variables are statistically significant predictors, or producing a simpler model, the p-value approach is preferred.
- Regardless of the approach used, model conditions must still be verified.

Adjusted R^2 vs. p-value approaches

The two approaches are similar, but they sometimes lead to different models, with the adjusted R^2 approach tending to include more predictors in the final model.

- When the sole goal is to improve prediction accuracy, use R^2. This is commonly the case in machine learning applications.
- When we care about understanding which variables are statistically significant predictors, or producing a simpler model, the p-value approach is preferred.
- Regardless of the approach used, model conditions must still be verified.

Better: monitor for multicollinearity.

Variance inflation factor (VIF)

For a linear model $y \sim x_1, x_2, \ldots, x_k$, we define ith **variance inflation factor (VIF)** as

$$\text{VIF}_i = 1/(1 - R_i^2)$$

where R_i is the R^2 for the linear model $x_i \sim x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$.

Variance inflation factor (VIF)

For a linear model $y \sim x_1, x_2, \ldots, x_k$, we define ith **variance inflation factor (VIF)** as

$$\mathrm{VIF}_i = 1/(1 - R_i^2)$$

where R_i is the R^2 for the linear model $x_i \sim x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$.

Multicollinearity occurs when two or more explanatory variables have a high coefficient of determination $R^2 \approx 1$,

Variance inflation factor (VIF)

For a linear model $y \sim x_1, x_2, \ldots, x_k$, we define ith **variance inflation factor (VIF)** as

$$\mathrm{VIF}_i = 1/(1 - R_i^2)$$

where R_i is the R^2 for the linear model $x_i \sim x_1, x_2, \ldots, x_{i-1}, x_{i+1}, \ldots, x_k$.

Multicollinearity occurs when two or more explanatory variables have a high coefficient of determination $R^2 \approx 1$, or equivalently, some VIF_i is big (common threshold: $\mathrm{VIF}_i > 5$).

In SAS: check for pairwise multicollinearity

Check with pairwise scatterplot:

```
PROC SGSCATTER DATA=BMIREG;
 TITLE "Scatterplot MATRIX for BMI DATA";
 MATRIX bmi age probl chobl;
RUN;
```

In SAS: check for pairwise multicollinearity

Check with pairwise correlations:

```
PROC CORR DATA=BMIREG;
 Var bmi age probl chobl alcbl;
RUN;
```

In SAS: check for pairwise multicollinearity

Pearson Correlation Coefficients, N = 100 Prob > |r| under H0: Rho=0				
	bmi	**age**	**probl**	**chobl**
bmi	1	-0.19965	0.05354	-0.04957
BMI at Baseline		0.0464	0.5968	0.6243
age	-0.19965	1	0.1925	0.14645
Age at Baseline	0.0464		0.055	0.146
probl	0.05354	0.1925	1	-0.29209
Protein (%kcal) at Baseline	0.5968	0.055		0.0032
chobl	-0.04957	0.14645	**-0.29209**	1
Dietary Carbohydrates	0.6243	0.146	**0.0032**	

317

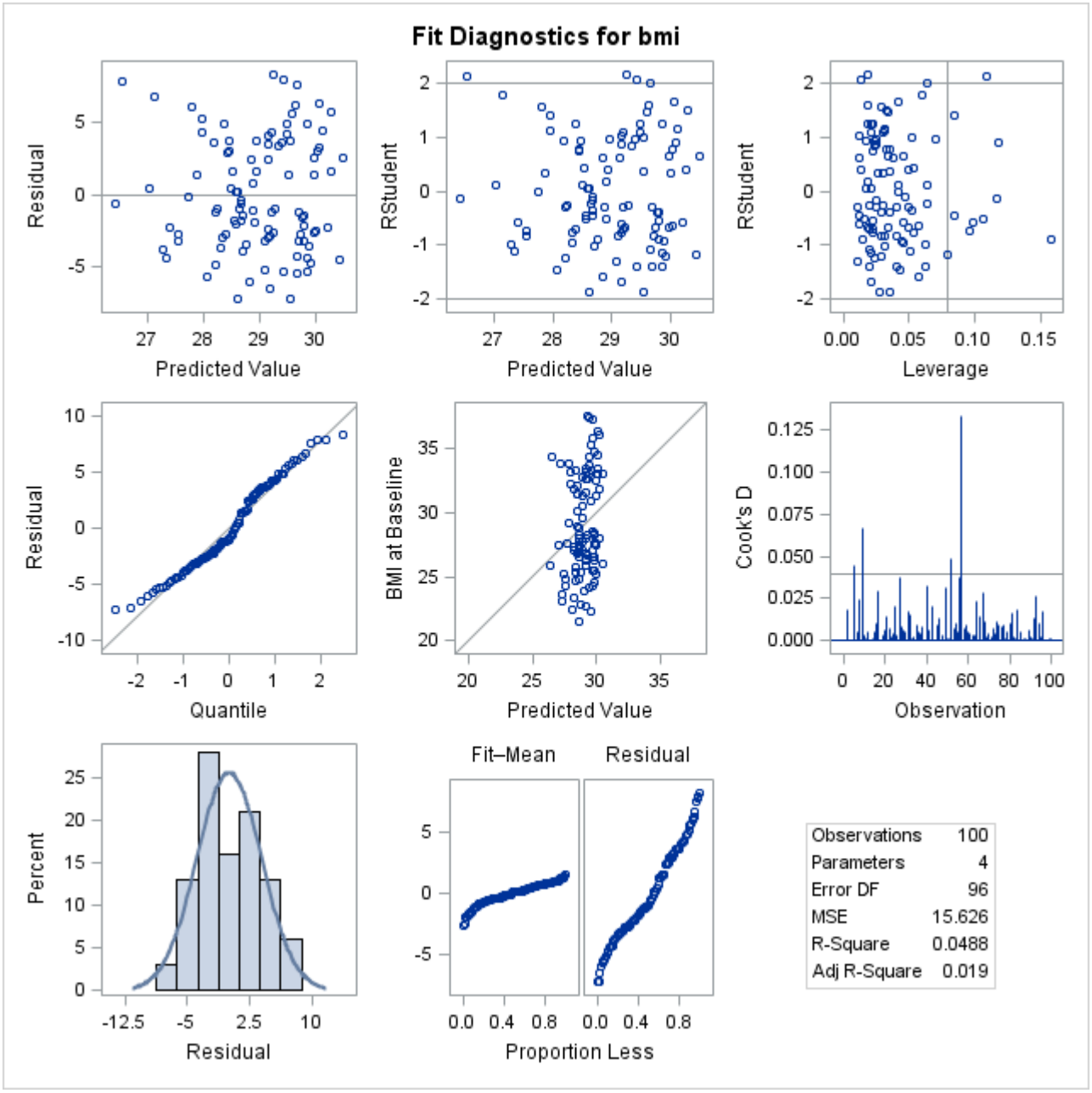
High Correlation potential problem in Regression - (multi) collinearity

In SAS: check for pairwise multicollinearity

Options:

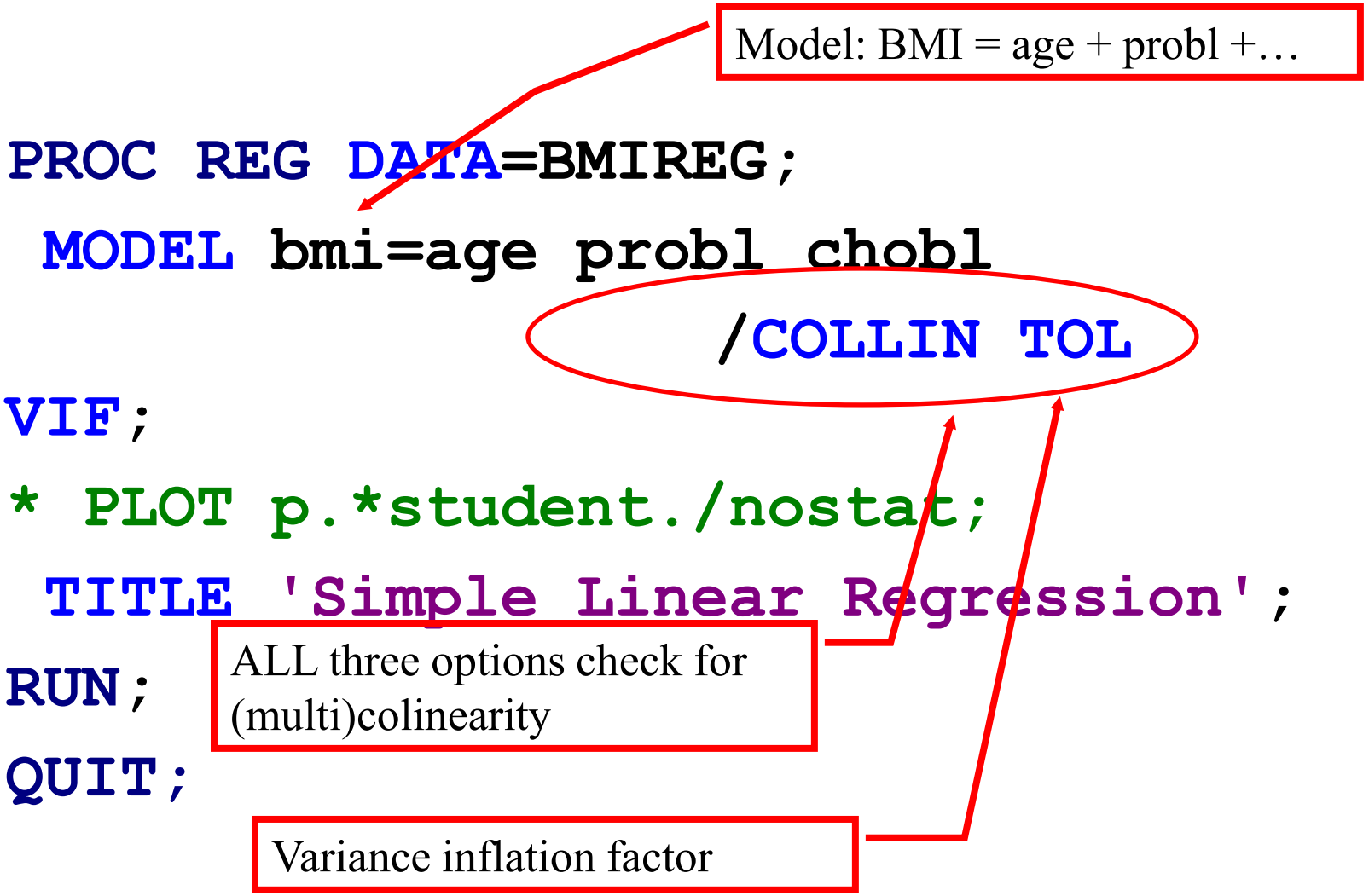
SELECTION=FORWARD, SELECTION=BACKWARD are p-value based.

SELECTION=RSQUARE and SELECTION=ADJRSQ gives all possible subsets of explanatory models used as linear models.

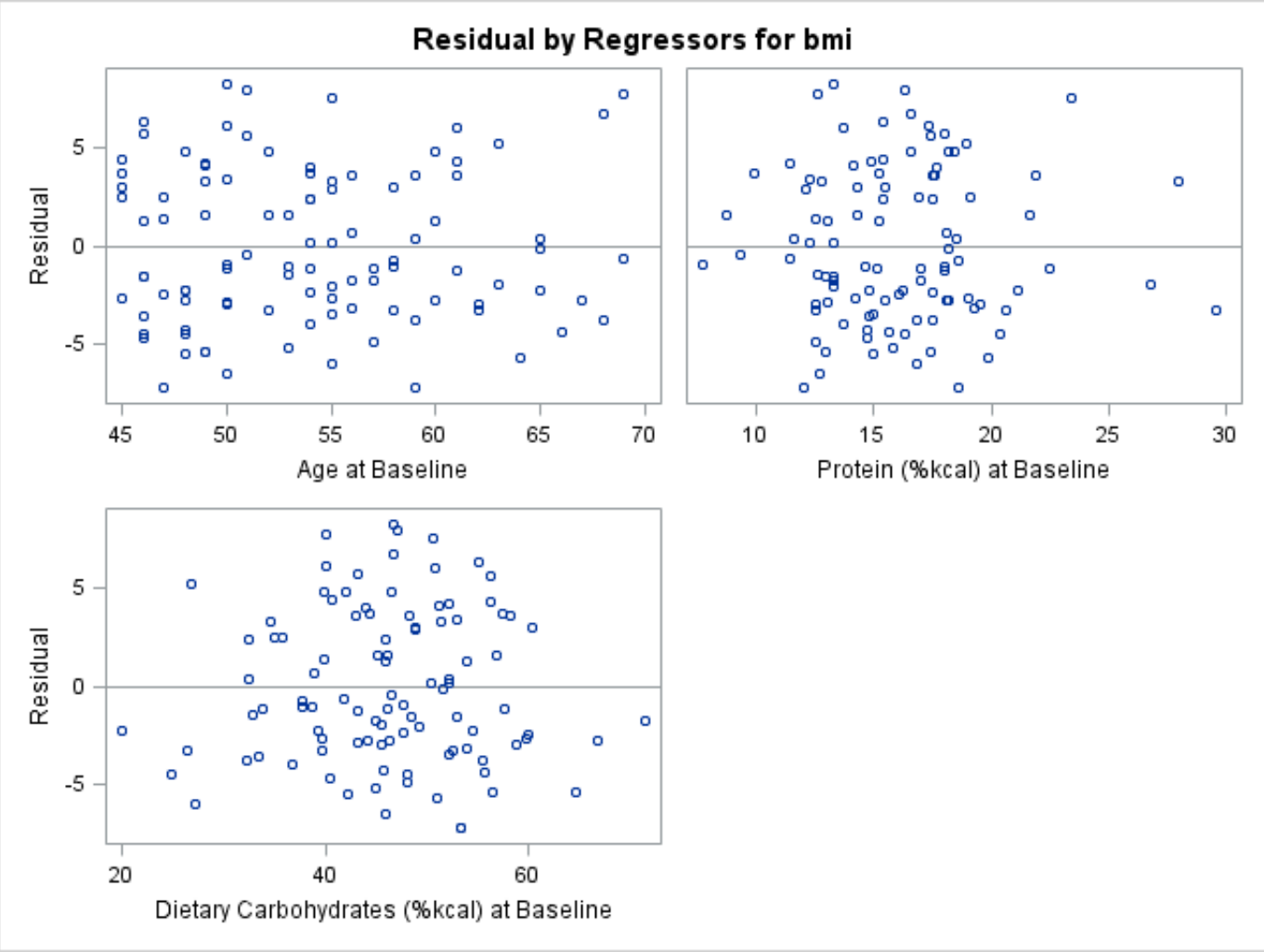


Diagnostic Plots

MLR: Fitting A Model



Diagnostic Plots: Residual Plots BY Independent Var



Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	3	76.8994	25.6331	1.64	0.1851
Error	96	1500.1	15.626		
Corrected Total	99	1577			

Root MSE	3.95298	R-Square	0.0488
Dependent Mean	28.9808	Adj R-Sq	0.019
Coeff Var	13.64		

R-Square close to 1 -> Model **Practically** useful

Parameter Estimates								
Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	Intercept	1	34.4093	4.05069	8.49	<.0001	.	0
age	Age at Baseline	1	-0.1366	0.06433	-2.12	0.0364	0.91803	1.08928
probl	Protein (%kcal) at Baseline	1	0.1072	0.11587	0.93	0.3572	0.85811	1.16535
chobl	Dietary Carbohydrates (%kcal) at Baseline	1	0.00506	0.04586	0.11	0.9123	0.87202	1.14676

P-values for each factor in model. If p-value small, factor important/has an effect on dependent variable

Parameter Estimates

Variable	Label	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Tolerance	Variance Inflation
Intercept	Intercept	1	34.4093	4.05069	8.49	<.0001	.	0
age	Age at Baseline	1	-0.1366	0.06433	-2.12	0.0364	0.91803	1.08928
probl	Protein (%kcal) at Baseline	1	0.1072	0.11587	0.93	0.3572	0.85811	1.16535
chobl	Dietary Carbohydrates (%kcal) at Baseline	1	0.00506	0.04586	0.11	0.9123	0.87202	1.14676

If Tolerance very small (<.2) for a factor, problem with collinearity

If inflation very large (>5) for a factor, problem with collinearity

If proportions large (>.5) in at least two factors in a row, problem with collinearity

Collinearity Diagnostics

Number	Eigenvalue	Condition Index	Proportion of Variation			
			Intercept	age	probl	chobl
1	3.92092	1	0.00062	0.00082	0.00269	0.00211
2	0.0573	8.27203	0.00044	0.00022	0.41827	0.25557
3	0.01526	16.0269	0.06593	0.34223	0.52737	0.6496

Model Selection: Who stays and who goes?

Methods for selecting which variables to keep in model:

1.Backward Selection

Put all variables in model and staring sequentially removing variables with largest p-value. Stop when all p-values < alpha (.05)

1.Forward selection

Put variable with largest correlation in model

Fit all models with 2 variables and select the one with the smallest p-value.

Stop when all p-values for variables not in model > alpha (.05)

1.Stepwise

Combination of the two above.

Note: Not all methods necessarily will result in the same model

Model Selection: Who stays and who goes?

```
/* Variable SELECTION: Backward   */
PROC REG DATA=BMIREG;
 MODEL bmi=age probl chobl/SELECTION=backward;
  TITLE 'Simple Linear Regression: Backward
SELECTION';
RUN;

/* Variable SELECTION:  forward  */
PROC REG DATA=BMIREG;
 MODEL bmi=age probl chobl/SELECTION=forward;
  TITLE 'Simple Linear Regression: forward
SELECTION';
RUN;

/* Variable SELECTION: Stepwise   */
PROC REG DATA=BMIREG;
 MODEL bmi=age probl chobl/SELECTION=Stepwise
                    SLENTRY = 0.10 SLSTAY = 0.05;
 TITLE 'Simple Linear Regression: Stepwise
```

SLE < = SLS to avoid infinite loops

Variable Selection: Backward

Backward Elimination: Step 2

Variable probl Removed: R-Square = 0.0399 and C(p) = 0.8986

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	62.8575	62.8575	4.07	0.0464

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	35.66553	3.33738	1764.5141	114.21	<.0001
age	-0.12363	0.06129	62.85752	4.07	**0.0464**

Summary of Backward Elimination

Step	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	**chobl**	Dietary Carbohydrates (%kcal) at Baseline	2	0.0001	0.0486	2.0122	0.01	0.9123
2	**probl**	Protein (%kcal) at	1	0.0088	0.0399	0.8986	0.9	0.3463

Variable Selection: Forward

Analysis of Variance					
Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	2	76.7089	38.3545	2.48	0.0891
Error	97	1500.29	15.4669		
Corrected Total	99	1577			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	34.62815	3.51449	1501.5358	97.08	<.0001
age	-0.13502	0.0625	72.18924	4.67	0.0332
probl	0.10298	0.10882	13.85142	0.9	0.3463

Summary of Forward Selection								
Step	Variable Entered	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	age	Age at Baseline	1	0.0399	0.0399	0.8986	4.07	0.0464
2	probl	Protein (%kcal) at Baseline	2	0.0088	0.0486	2.0122	0.9	0.3463

Variable Selection: Stepwise

Stepwise Selection: Step 1

Analysis of Variance

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	1	62.8575	62.8575	4.07	0.0464
Error	98	1514.14	15.4504		
Corrected Total	99	1577			

Variable	Parameter Estimate	Standard Error	Type II SS	F Value	Pr > F
Intercept	35.66553	3.33738	1764.5141	114.21	<.0001
age	-0.12363	0.06129	62.85752	4.07	0.0464

Summary of Stepwise Selection

Step	Variable Entered	Variable Removed	Label	Number Vars In	Partial R-Square	Model R-Square	C(p)	F Value	Pr > F
1	age		Age at Baseline	1	0.0399	0.0399	0.898	4.07	0.0464