EDUCATION AND VOCABULARY

MULTIPLE REGRESSION IN ACTION

EDUCATION AND VOCABULARY

- 5-10 hours of input weekly is enough to pick up a new language (Schiff & Myers, 1988).
- Dutch children spend 5.5 hours/day in front of a screen (Valkenburg, 2013).
- Most of this input is in English.
- How much does education contribute?

RESEARCH QUESTION

Does the amount of time children are taught English weekly predict the size of their English vocabulary, or are there other factors – and if so, to what extent are they correlated with English vocabulary?

STUDY

- Participants
  - 72 Dutch children;
  - Primary school classes 5 and 6;
  - Age 8 – 10, but expressed in months (m=113.5);
  - 33 males, 39 females.

- Schools matched for
  - Low-risk;
  - High SES;
  - Urban environment;
  - No other official languages (like Frisian);
  - Cito scores.

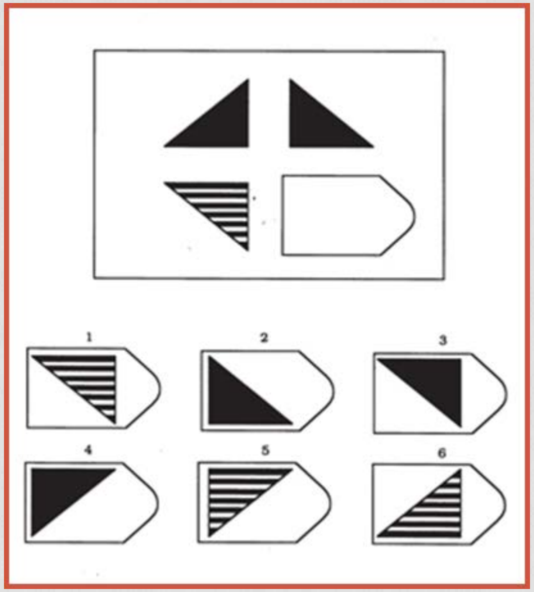
STUDY

- Hours of English:
  - School 1, which teaches 4 hours of English weekly. We tested 32 students, 4 of which were left out due to missing or unusable data*.
  - School 2, which teaches 2 hours of English weekly. We tested 34 students, 10 of which were left out .
  - School 3, which teaches no English in groups 5 and 6 (control). We tested 31 students, 11 of which were left out.

* Technical problems, learning disabilities, etc.

TOOLS

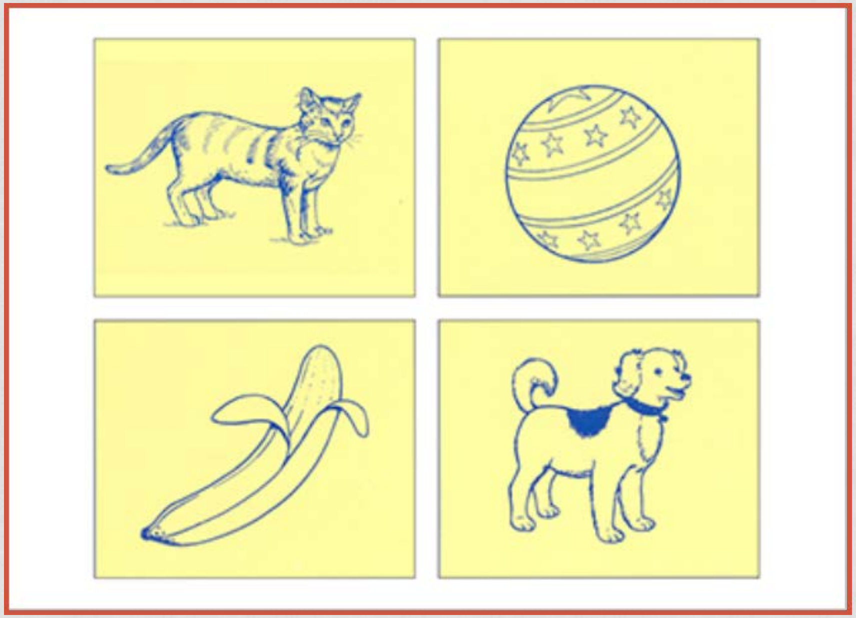
- Raven intelligence test, power version:
  - 48 questions;
  - 20 minutes;
  - Score = total correct.



Example Raven exercise, from http://www.talentlens.nl

TOOLS

- Peabody NL (language aptitude):
  - Dutch words presented over headphone;
  - Subjects must click on matching picture out of 4;
  - Score = total correct;
  - Increasing difficulty;
  - Max score = 204.



Example Peabody NL exercise. Test developed by Pearson and software developed by Dr. Claire Stevenson, University of Leiden.

TOOLS

- Peabody EN (English vocabulary):
  - English words presented over headphone;
  - Subjects must click on matching picture out of 4;
  - Score = total correct;
  - Increasing difficulty;
  - Max score = 228.

Example Peabody EN exercise. Test developed by Pearson and software developed by Dr. Claire Stevenson, University of Leiden.

FORMULA

Peabody EN score$_i$ = (b_0 + b_1 hours$_i$ + b_2 aptitude$_i$ + b_3 age$_i$ + b_4 intelligence$_i$) + ε_i

SIMPLE REGRESSION

R Output

```
> englishSR<-lm(pben ~ hours, data=english)
> summary(englishSR)

Call:
lm(formula = pben ~ hours, data = english)

Residuals:
   Min     1Q Median     3Q    Max
-36.87 -25.34 -15.32  20.57 110.91

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)   31.318      6.886   4.548 2.21e-05 ***
hours          4.388      2.505   1.752   0.0842 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 34.39 on 70 degrees of freedom
Multiple R-squared:  0.04199,   Adjusted R-squared:  0.0283
F-statistic: 3.068 on 1 and 70 DF,  p-value: 0.08424
```

Interpretation

- Hours of English explains only 4.2% of the variation in PBEN.
- Not significant.

MULTIPLE REGRESSION

R Output

```
> englishMR<-lm(pben ~ hours + age + raven + pbnl, data=english)
> summary(englishMR)

Call:
lm(formula = pben ~ hours + age + raven + pbnl, data = english)

Residuals:
    Min      1Q  Median      3Q     Max
-46.274 -15.792  -3.031  18.155  58.196

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) -275.7125    46.2748  -5.958 1.05e-07 ***
hours         -0.3710     2.2422  -0.165 0.869098
age            1.2612     0.3471   3.633 0.000543 ***
raven          1.2722     0.4780   2.661 0.009732 **
pbnl           1.4268     0.2486   5.739 2.51e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 25.55 on 67 degrees of freedom
Multiple R-squared:  0.4939,    Adjusted R-squared:  0.4637
F-statistic: 16.34 on 4 and 67 DF,  p-value: 2.172e-09
```

Interpretation

- Age, intelligence and aptitude account for an extra 45%.
- Adjusted R^2 is 3% less.
- Highly significant at $P < 0.001$.

INTERPRETATION

- As hours increases by one unit, PBEN decreases by 0.37 units (!)
  - However, the contribution of this variable to the model is highly insignificant at $P = 0.87$.
- As age increases by one unit, PBEN increases by 1.25 units.
  - Highly significant contribution at $P < 0.001$
- As intelligence increases by one unit, PBEN increases by 1.19 units.
  - Highly significant contribution at $P < 0.01$
- As aptitude increases by one unit, PBEN increases by 1.5 units.
  - Highly significant contribution at $P < 0.001$

STANDARDIZED B-VALUES

R Output

```
> lm.beta(englishMR)
hours        age         raven       pbnl
-0.01732222  0.31904493  0.27639488  0.51697292
```

Interpretation

- Number of SDs by which PBEN will change as each of the predictors changes by 1 SD (all other predictors being equal!).
- Directly comparable;
- Better insight into weight of each variable.

CONFIDENCE INTERVALS

R Output

> confint(englishMR)

	2.5 %	97.5 %
(Intercept)	-368.0773784	-183.347680
hours	-4.8464871	4.104587
age	0.5683334	1.954122
raven	0.3180700	2.226378
pbnl	0.9305210	1.923057

Interpretation

- The confidence bands for each of the predictors is small, except for hours.
- Hours crosses 0: sometimes the relationship is positive, sometimes negative.
- BAD.

COMPARING MODELS

R Output

```
> anova(englishSR, englishMR)
Analysis of Variance Table

Model 1: pben ~ hours
Model 2: pben ~ hours + age + raven + pbnl
  Res.Df   RSS  Df  Sum of Sq      F     Pr(>F)
1     70 82790
2     67 43739   3      39051  19.94  2.401e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Interpretation

- EnglishMR is a signficantly better fit to the data compared to EnglishSR, $F(3, 67) = 19.94$, $p < 0.001$.

DIAGNOSTICS

R Output

```
> english$standardized.residuals<-rstandard(englishMR)
> english$large.residual<-english$standardized.residuals > 2 |
english$standardized.residuals < -2
sum(english$large.residual)
[1] 3

> english[english$large.residual, c("pben", "age", "raven", "pbnl", "hours",
"standardized.residuals")]
    pben   age   raven   pbnl   hours   standardized.residuals
1   149    128   39      110    4       2.389620
48  151    117   41      121    2       2.285620
56  92     109   27      99     0       2.198725
```

Interpretation

- Sample = 72
- 95% of residuals should be within +/- 2 (SD).
- 5% should be outside.
- 5% of 72 = 3.6
- 3 or 4 outliers
- We have 3.
- Fine.

DIAGNOSTICS

R Output

```
> english$cooks<-cooks.distance(englishMR)
> english$leverage<-hatvalues(englishMR)
> english$covariance<-covratio(englishMR)
> english[english$large.residual, c("cooks", "leverage", "covariance")]
```

	cooks	leverage	covariance
1	0.11501253	0.09149260	0.7601336
48	0.12934210	0.11015771	0.8073542
56	0.05533664	0.05413405	0.7837935

Interpretation

- Cook's distance should be < 1.
- Leverage should be < 2(k + 1/n);
  - 2(5/72) = 0.14
- Covariance ratio
  - $CVR_i < 1 + [3(k + 1)/n]$
  - $CVR_i < 1 + [3(4 + 1)/72] = 1.08$
  - $CVR_i > 1 - [3(k + 1)/n]$
  - $CVR_i > 1 - [3(4 + 1)/72] = 0.79$
- #1 is lowish, but see Cook's distance.

INDEPENDENCE

R Output

```
> dwt(englishMR)
 lag Autocorrelation D-W Statistic p-value
   1      0.07124528      1.778073   0.228
 Alternative hypothesis: rho != 0
```

Interpretation

- Durbin-Watson tests assumption of independent errors.
- Should be close to 2 and not <1 or >3.
- Fine at 1.78.

NO MULTICOLLINEARITY

R Output

```
> vif(englishMR)
hours       age         raven       pbnl
1.451289    1.020795    1.427768    1.074327

> 1/vif(englishMR)
hours       age         raven       pbnl
0.6890425   0.9796286   0.7003941   0.9308155

> mean(vif(englishMR))
[1] 1.243545
```

Interpretation

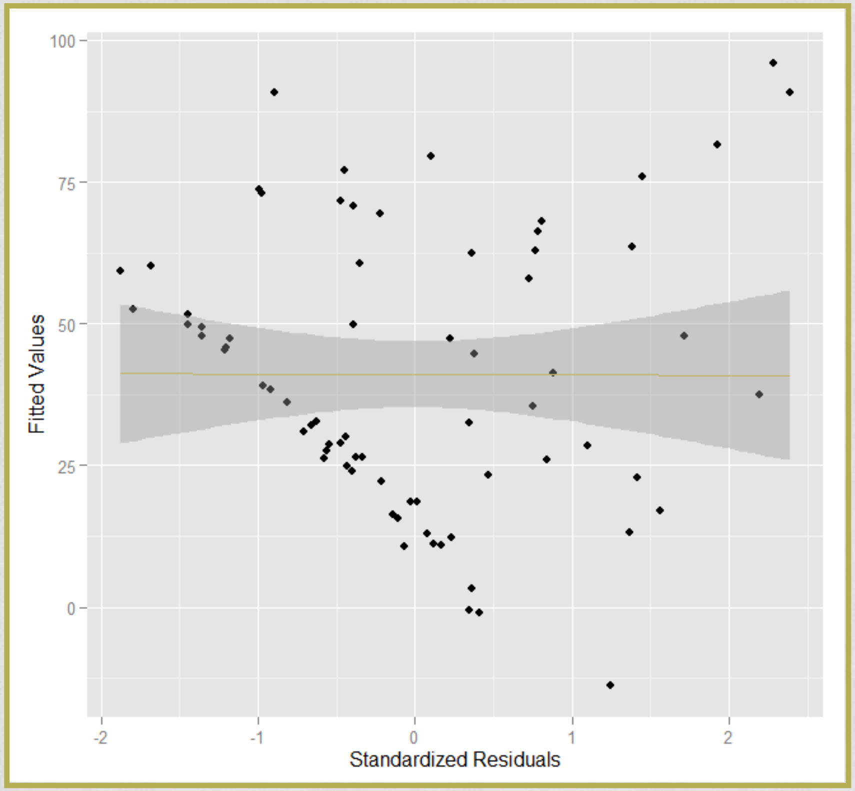
- VIF to assess multicollinearity.
- Tolerance = 1/VIF.
- Largest VIF > 10 means problem.
- Mean VIF much > 1 means problem.
- Tolerance < 0.2 means potential problem.
- All fine.

RESIDUALS

R Output

```
> english$fitted <- englishMR$fitted.values
> scatterResiduals<-ggplot(english, aes(standardized.residuals, fitted))
> scatterResiduals<-scatterResiduals + geom_point() + geom_smooth(method="lm", colour="darkkhaki") + labs(x="Standardized Residuals", y="Fitted Values")
> scatterResiduals
```

Visualizing residuals

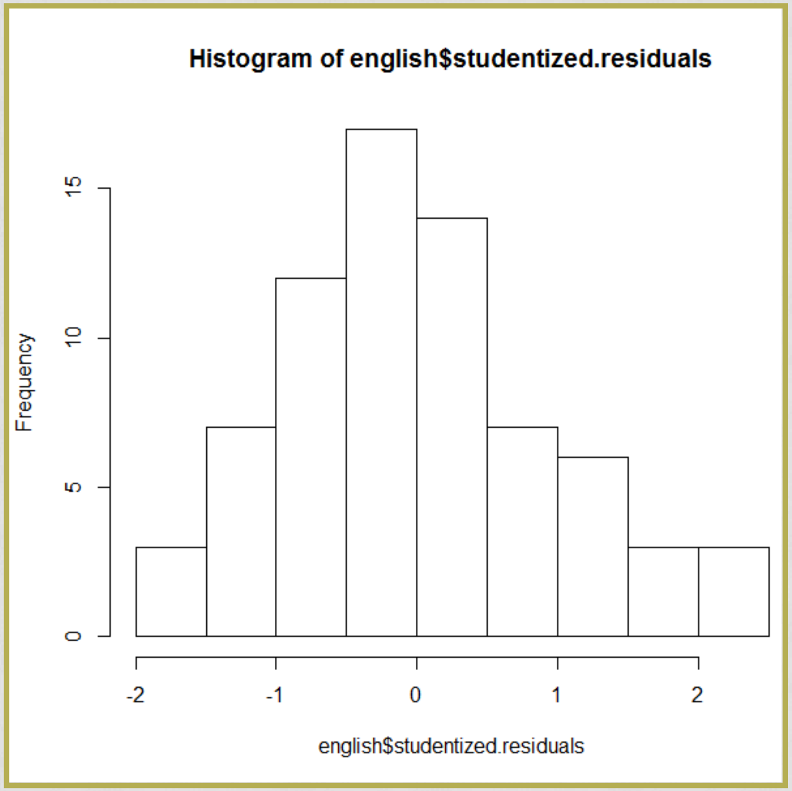


RESIDUALS

R Output

> hist(english$studentized.residuals)

Visualizing residuals



INTERPRETING RESIDUALS

- Some heteroscedascity and non-linearity.
- Distribution of residuals seems normal.

CONCLUSION

- Assumption of homoscedascity and linearity of residuals violated.
- Findings cannot be generalized beyond sample (yet).
- Options:
  - Logistic regression
  - Robust regression

CONCLUSION

- Hours of education does not predict PBEN score.
- Rather, a combination of age, intelligence and language aptitude does.

REFERENCES

- Field, A. et al (2012). *Discovering statistics using R.* London: Sage Publications Ltd.
- Moore, D. S. et al (2012). *Introduction to the practice of statistics.* New York: W. H. Freeman and Company.
- Schiff-Myers, N., Klein, H. (1985). Some phonological characteristics of the speech of normal-hearing children of Deaf parents. *Journal of Speech and Hearing Research, 28*(4), 466-474.
- Valkenburg, P. et al. (2013). Developing and validating the perceived parental media mediation scale: A self-determination perspective. *Human Communication Research, 39.* 445-469.