

Part 4
EXPLAINED VARIANCE
PREDICTION ERROR

effect size

significance = effect.size * sample.size
(Rosnow & Rosenthal 2008)

measures of effect size:

- *t*-test $D=M1-M2$; $d=D/s$
- ANOVA η^2 , $\text{part.}\eta^2$, ω
- regression R^2

effect size measures

- Cohen $d = (M1-M2)/s = 2t/\sqrt{df}$
- point-biserial $r = t / \sqrt{t^2+df}$
correlation between dichotomous predictor and outcome variable
where $df=N-2$
Rosenthal, Rosnow & Rubin, Formula 2.3

R² measure

- normal regression:
 R^2 is proportion of (random) variance explained by (fixed) predictors
- mixed-effects regression:
proportion of *which* random variance?
– schools/students/tests
– speakers/phrases
– (listeners & items) / words

R² measure

- R^2 is regarded as
“proportional reduction in prediction error”
(Snijders & Bosker, 1999:102; Luke, 2004:35)
- additional predictor reduces random variability
- compare model of interest with suitable null model

Finnish F0 peak timing

- data kindly provided by A. Arnhold
- DV is **HdistM1end*1000**,
distance from end of mora1, in **ms**
- DV is not normally distributed (*qqnorm*)
- subjects and items as crossed random effects

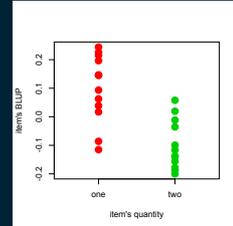
empty model

```
> summary(arnholdm.m00)
Random effects:
Groups   Name      Variance Std.Dev.
item     (Intercept)  104.00  10.198
subject  (Intercept)  220.14  14.837
Residual                    10148.14 100.738
Number of obs: 2367, groups: item, 24; subject, 17

Fixed effects:
              Estimate Std. Error t value
(Intercept)  -15.512      4.652   -3.335
```

inspecting BLUPs

- Best Linear Unbiased Predictor
- random effect of item
- while taking into account (no) fixed effects and subjects' effects
- broken down by item's quantity



quantity predictor

- vowel quantity may affect DV
- this predictor would explain between-item variance, not between-subject or residual variance

```
> summary(arnholdm.m01)
Linear mixed model fit by maximum likelihood
Formula: HdistM1end * 1000 ~ (1 | subject) + (1 | item) + quantity
Random effects:
Groups   Name      Variance Std.Dev.
item     (Intercept)  72.419  8.5099
subject  (Intercept)  219.070 14.8010
Residual                    10146.757 100.7311
Number of obs: 2367, groups: item, 24; subject, 17
```

```
> summary(arnholdm.m01)
Linear mixed model fit by maximum likelihood
Formula: HdistM1end * 1000 ~ (1 | subject) + (1 | item) + quantity
...
Fixed effects:
              Estimate Std. Error t value
(Intercept)  -21.355      5.275   -4.049
quantitytwo   11.566      5.419    2.134
```

R²

	m00	m01
var(items)	104	72
var(resid)=10148/97	105	105
sum	209	177

- $R^2(\text{items}) = 1 - (177/209) = 0.153$

R^2

- random variance is diminished by .15 (1/6th)
- random variance due to subjects is ignored
 - crossed, independent, orthogonal

Questions ?