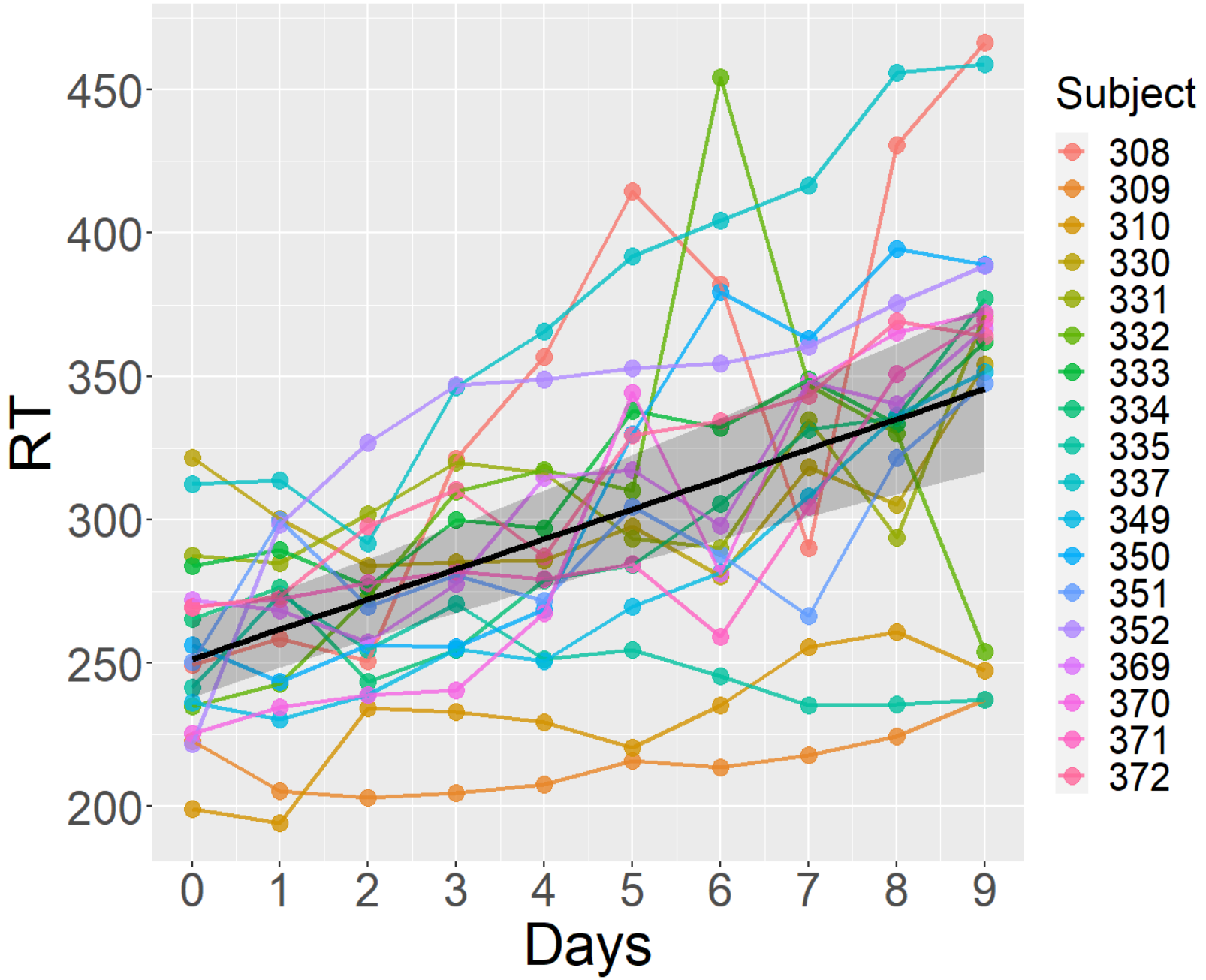
"A profitable way to build understanding and confidence is through data simulation. **If you can create data sets** by sampling from a population for which you **know the ground truth** about the population parameters you are interested in (e.g., mean and standard deviation of each group), **you can check how often and under what circumstances a statistical model will give you the correct answer**"

(DeBruine et al., 2021)

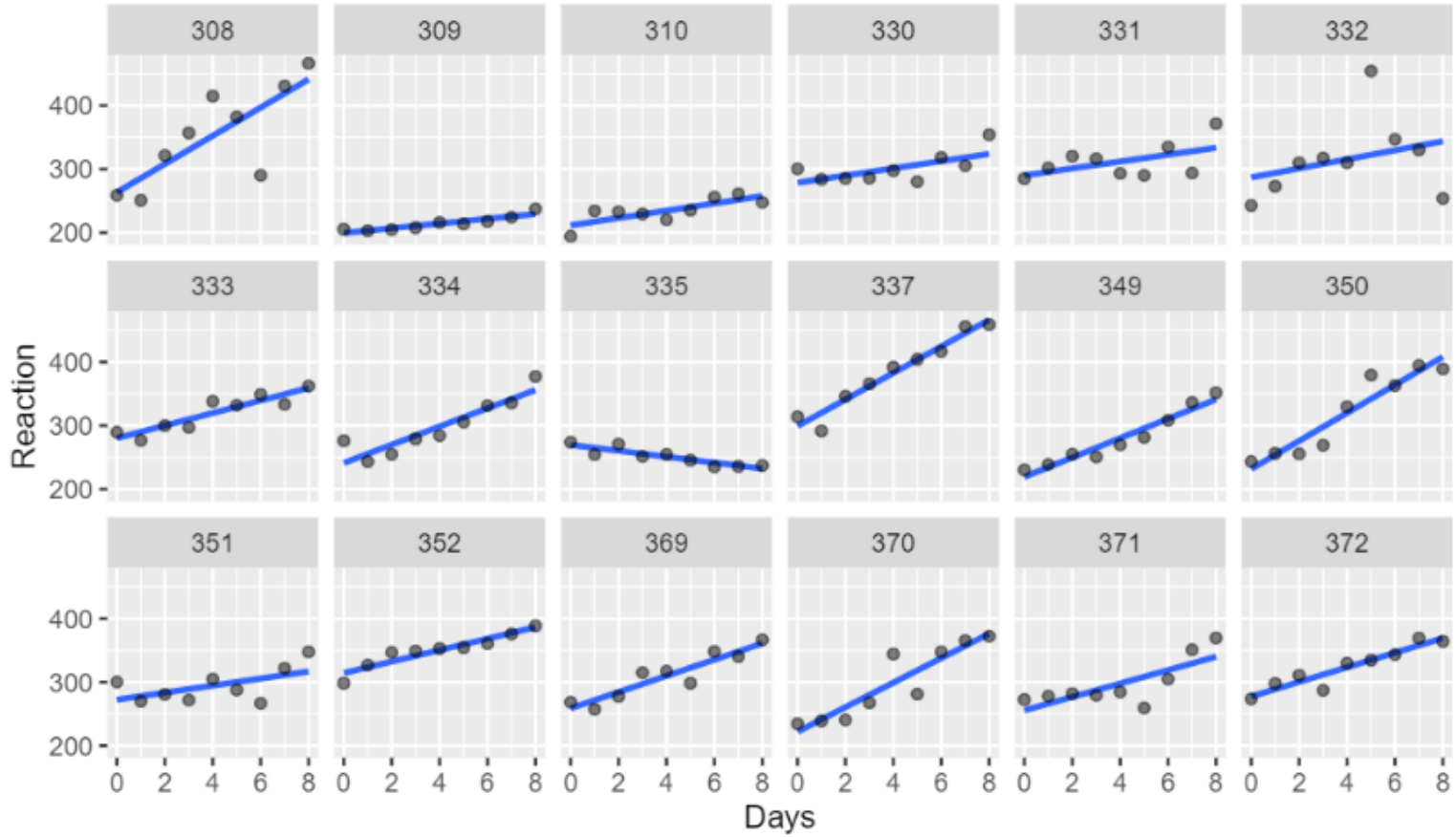
RECAP

```
> fit = lmer(Reaction ~ Days + (Days | subject), data = d)
> summary(fit)
Linear mixed model fit by REML. t-tests use Satterthwaite's
['lmerModLmerTest']
Formula: Reaction ~ Days + (Days | Subject)
   Data: d

REML criterion at convergence: 1743.6

Scaled residuals:
    Min      1Q  Median      3Q     Max
-3.9536 -0.4634  0.0231  0.4634  5.1793

Random effects:
 Groups   Name        Variance Std.Dev. Corr
 Subject  (Intercept) 612.10   24.741
          Days         35.07    5.922   0.07
 Residual             654.94   25.592
Number of obs: 180, groups:  Subject, 18

Fixed effects:
            Estimate Std. Error      df t value Pr(>|t|)
(Intercept)  251.405      6.825  17.000  36.838  < 2e-16 ***
Days          10.467      1.546  17.000   6.771 3.26e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’
```



RECAP

```
> ggplot(data=d, aes(x=Days, y=Reaction)) + facet_wrap(~Subject,ncol=6)
+ geom_smooth(method="lm",se=FALSE,formula="y~x") +
+ geom_point(alpha=.5)
```

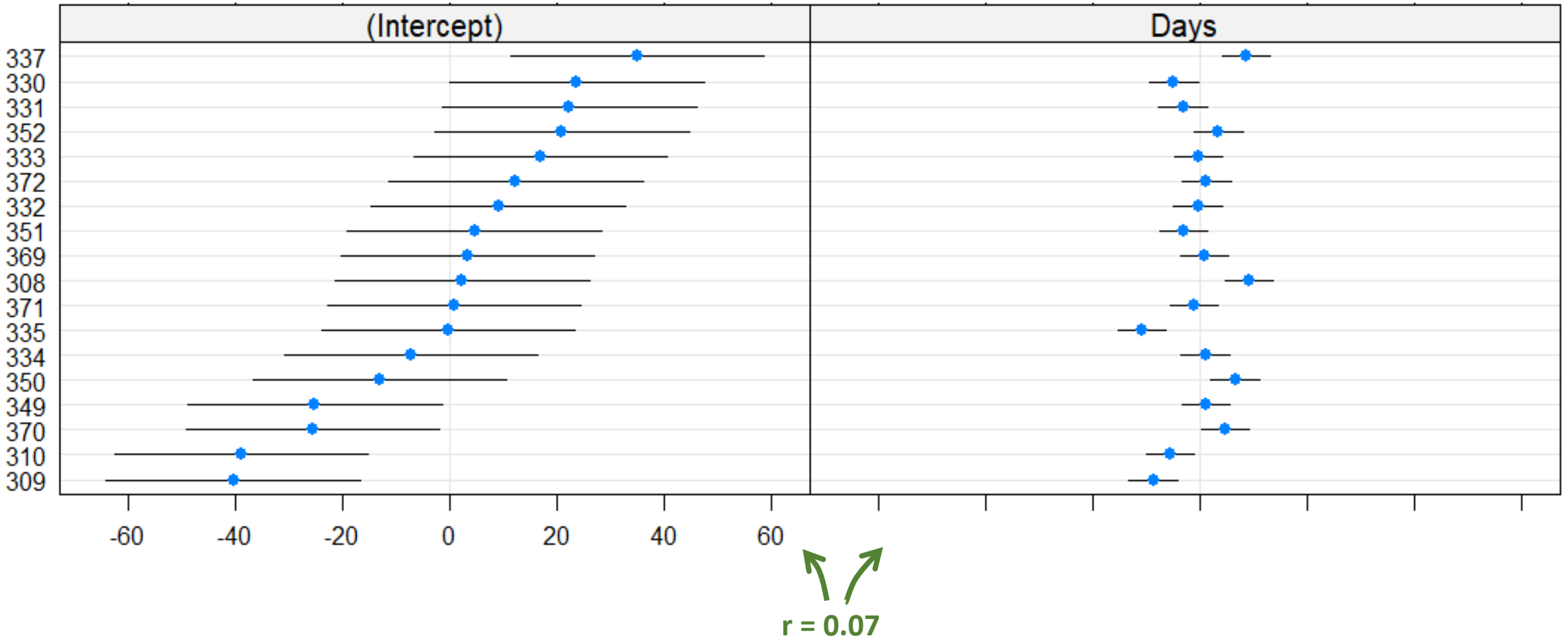


da slide del **prof. Altoè (*PsicoStat*)**: https://osf.io/b7tkp/

RECAP

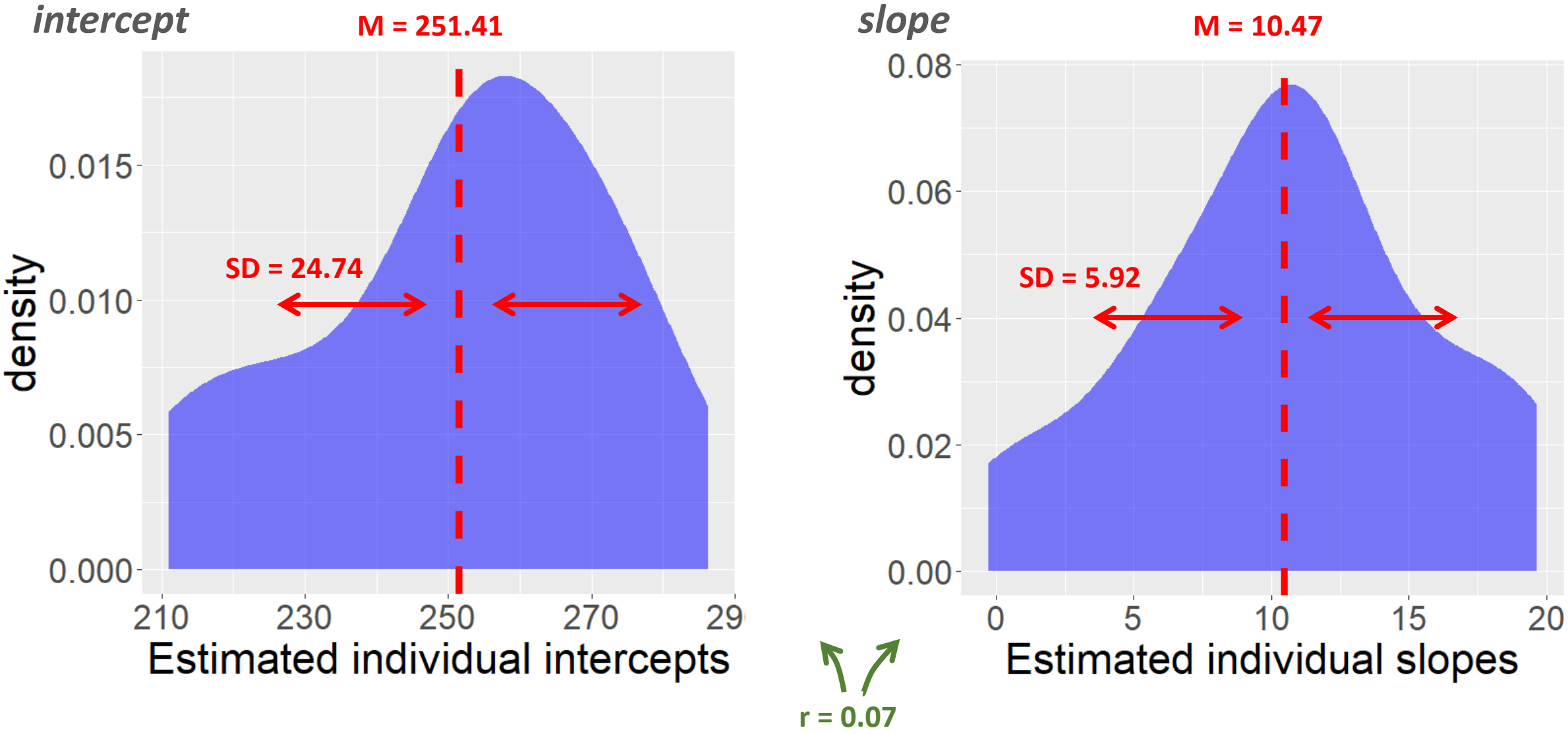
Distribuzione degli effetti random individuali

lattice :: dotplot(ranef(fit))



RECAP

Distribuzione degli effetti random individuali



Immaginiamo studio multilab come situazione analoga

- L'effetto random non è (per forza) il soggetto, ma il lab
- Invece di avere soggetti che hanno più giorni come misure ripetute, ho laboratori che hanno più soggetti come misure ripetute
- Invece di avere un effetto medio generale nel campione che varia di soggetto in soggetto, ho un effetto medio generale nel multilab study (cruciale per l'inferenza) che varia di lab in lab (comunque importante per capire l'eterogeneità e variabilità degli effetti nei contesti)
- **ESEMPIO: ho uno studio multilab che coinvolge *k* laboratori, ciascuno dei quali raccoglie *n* casi vs *n* controlli e li confronta calcolando una *Standardized Mean Difference* (e.g., Cohen's *d*) sulla variabile *y***

ESEMPIO: MULTILAB

k = 10 laboratori, ciascuno raccoglie *n* = 100 casi clinici vs *n* = 100 controlli (N totale = 2000)

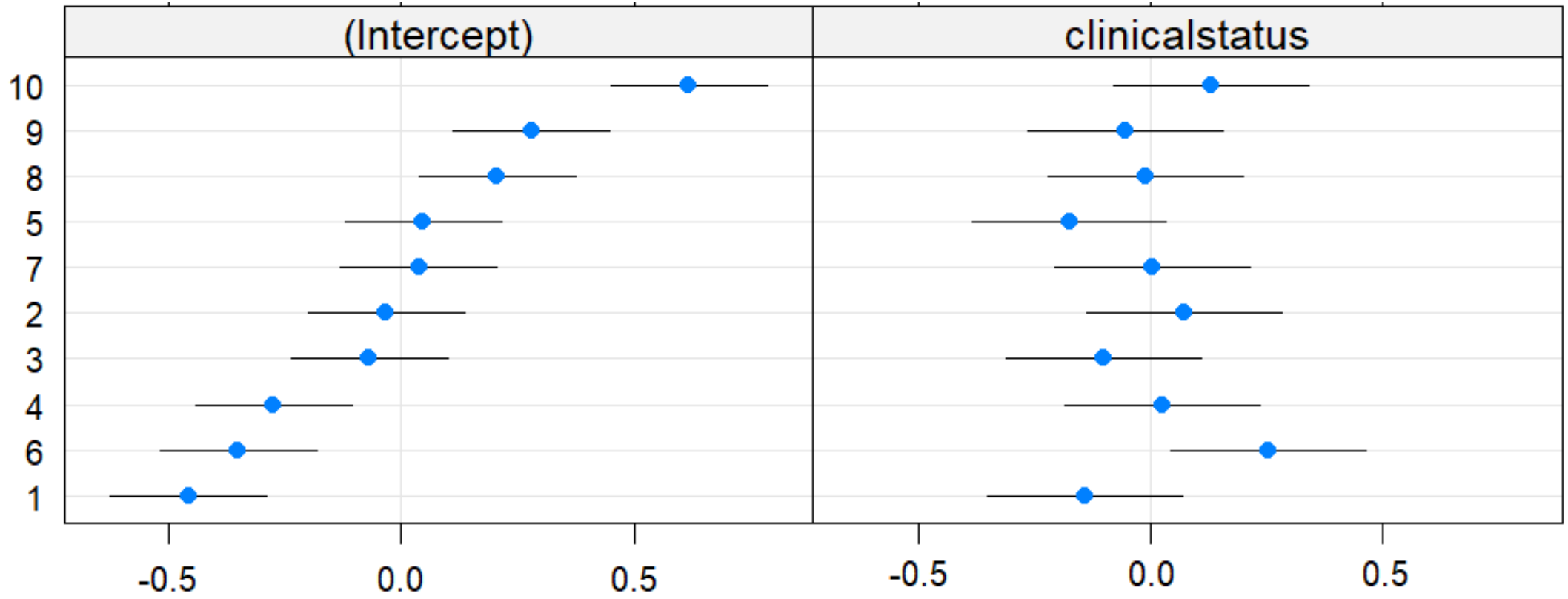
Parametri «*ground truth*» della simulazione: intercetta fissa = 0.0; slope fissa = -0.5;
SD intercetta random = 0.3; SD slope random = 0.2; SD residui = 1.0

```
> fit = lmer(y ~ clinicalstatus + (clinicalstatus | lab), data=d)
> summary(fit)
Random effects:
  Groups   Name           Variance Std.Dev. Corr
 lab      (Intercept)    0.10956  0.3310
          clinicalstatus 0.02807  0.1676   -0.04
 Residual                0.98949  0.9947
Number of obs: 2000, groups:  lab, 10

Fixed effects:
             Estimate Std. Error       df t value Pr(>|t|)
(Intercept)   0.05426    0.10930  8.99998   0.496    0.631
clinicalstatus -0.52172    0.06918  8.99996  -7.541 3.54e-05 ***
```

ESEMPIO: MULTILAB

k = 10 laboratori, ciascuno raccoglie *n* = 100 casi clinici vs *n* = 100 controlli (N totale = 2000)

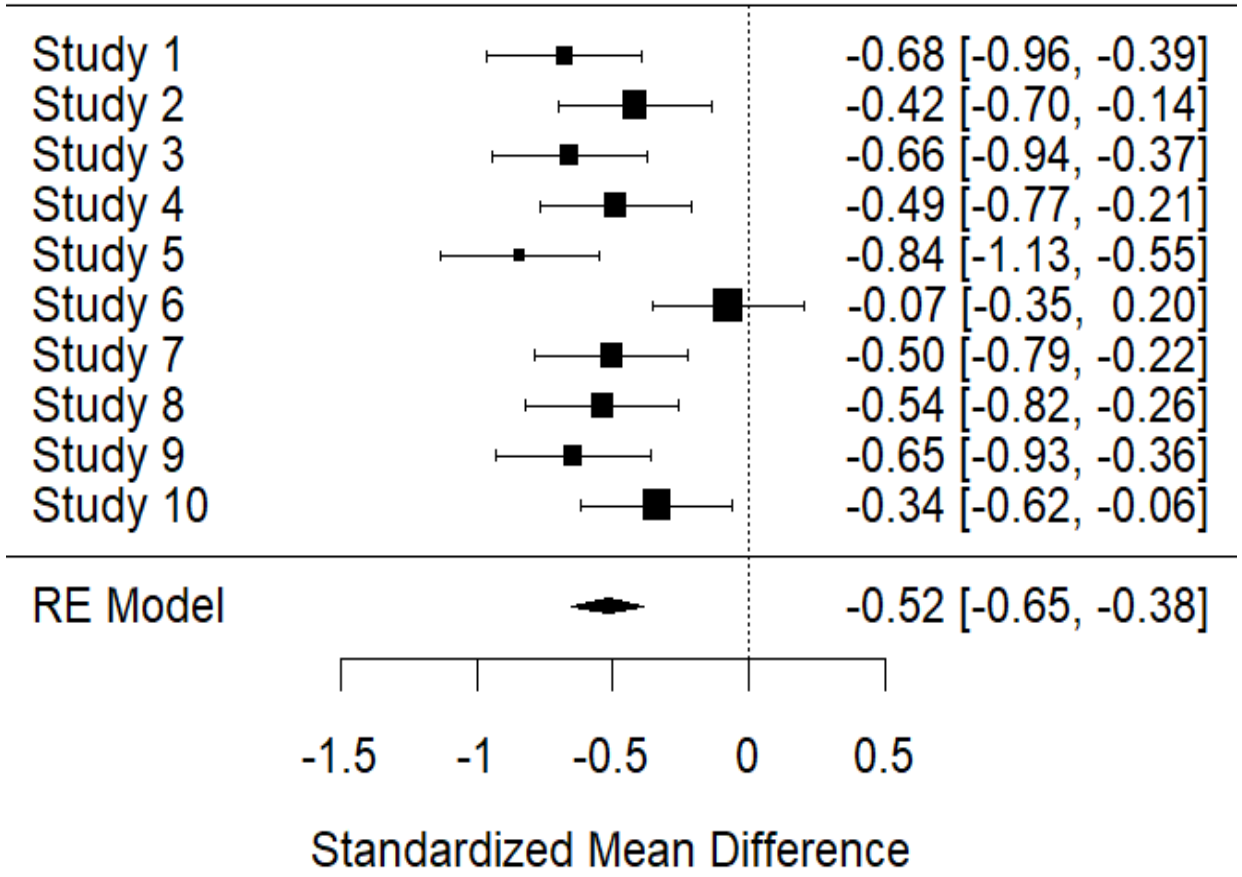
Parametri «*ground truth*» della simulazione: intercetta fissa = 0.0; slope fissa = -0.5;
SD intercetta random = 0.3; SD slope random = 0.2; SD residui = 1.0



ESEMPIO: MULTILAB

Rianalizzato come una meta-analisi a effetti random

```
> fitMeta = rma(yi,vi,data=d_agg)
> fitMeta
Random-Effects Model (k = 10; tau^2 estimator: REML)

tau^2 (estimated amount of total heter.): 0.0247 (SE=0.0214)
tau (square root of estimated tau^2 value):      0.1571
I^2 (total heterogeneity / total variability):   54.31%
H^2 (total variability / sampling variability):  2.19

Test for Heterogeneity:
Q(df = 9) = 19.7086, p-val = 0.0198

Model Results:

estimate      se     zval    pval    ci.lb    ci.ub
 -0.5161  0.0674  -7.6559  <.0001  -0.6482  -0.3840  ***
```



DALLA FORMULA ALLA SIMULAZIONE

$$y_{ij} = \beta_0 + \beta_1 * X_{ij} + T_{0j} + T_{1j} * X_{ij} + \varepsilon_{ij}$$

- y_{ij}: dato osservato nel soggetto i-esimo del j-esimo lab
- β_0: intercetta fissa (stimata per la popolazione)
- β_1: slope fissa (stimata per la popolazione)
- X_{ij}: valore del predittore X per l'i-esimo soggetto del j-esimo lab
- T_{0j}: intercetta random (stimata per il j-esimo lab)
- T_{1j}: slope random (stimata per il j-esimo lab)
- ε_{ij}: residuo (termine di errore per il soggetto i-esimo del j-esimo lab)

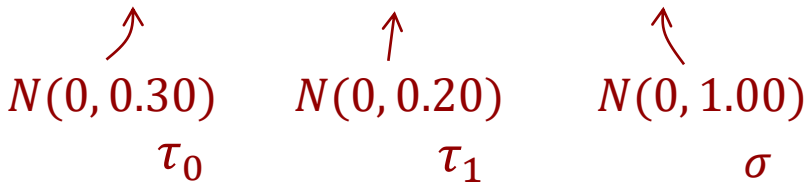
$$T_0 \sim N(0, \tau_0)$$
$$T_1 \sim N(0, \tau_1)$$
$$\varepsilon \sim N(0, \sigma)$$

Quando guardiamo > summary(fit) vediamo:
$\hat{\beta}_0, \hat{\beta}_1, \hat{\tau}_0, \hat{\tau}_1, \hat{\sigma}$

con > ranef(fit) vediamo anche:
$\hat{T}_{0j}, \hat{T}_{1j}$

DALLA FORMULA ALLA SIMULAZIONE

lab (j)	sub (i)	clinicalstatus (X_{ij})	β_0	β_1	T_{0j}	T_{1j}	ε_{ij}	y_{ij}
1	11	0	0.00	– 0.50	0.46	0.23	0.24	0.70
1	12	0	0.00	– 0.50	0.46	0.23	– 0.81	– 0.35
1	13	1	0.00	– 0.50	0.46	0.23	0.23	0.42
1	14	1	0.00	– 0.50	0.46	0.23	0.05	0.24
2	21	0	0.00	– 0.50	– 0.36	0.02	0.90	0.54
2	22	0	0.00	– 0.50	– 0.36	0.02	0.95	0.59
2	23	1	0.00	– 0.50	– 0.36	0.02	– 0.57	– 1.41
2	24	1	0.00	– 0.50	– 0.36	0.02	1.57	0.73
3	31	0	0.00	– 0.50	0.12	– 0.18	– 1.73	– 1.61
3	32	0	0.00	– 0.50	0.12	– 0.18	0.90	1.02
3	33	1	0.00	– 0.50	0.12	– 0.18	1.70	1.14
3	34	1	0.00	– 0.50	0.12	– 0.18	– 0.87	– 1.43



DALLA FORMULA ALLA SIMULAZIONE

```
> fit = lmer(y ~ clinicalstatus + (clinicalstatus | lab), data=d)
> summary(fit)
Random effects:
  Groups   Name           Variance Std.Dev. Corr
 lab      (Intercept)     0.09000  0.3000
          clinicalstatus  0.04000  0.2000   0.00
 Residual                 1.00000  1.0000
Number of obs: xxxxx, groups:  lab, xxxxx

Fixed effects:
             Estimate   Std. Error     df    t value  Pr(>|t|)
(Intercept)    0.00        xxxxx      xxx     xxxxx     xxxxx
clinicalstatus -0.50       xxxxx      xxx     xxxxx     xxxxx
```