MULTILEVEL 3 - Data set: EXAM

INTRODUZIONE

In questo dataset sono contenute 4059 osservazioni e le seguenti 9 variabili:

1. SCHOOL: id della scuola
2. NORMEXAM: score ottenuto all'esame normalizzato
3. SCHGEND: genere della scuola (mixed, boys, girls)
4. SCHAVG: intake score a livello di scuola
5. VR: verbal reasoning score a livello di studente
6. INTAKE: intake score a livello di studente
7. STANDLRT: LR test score
8. SEX: genere (M, F)
9. TYPE: tipologia di scuola (MXD, SNGL)
10. STUDENT: id dello studente

Analisi proposte:

1. Statistiche descrittive
2. Analisi multilevel

```r
#-- R CODE
library(car)
library(sjstats)
library(plotrix)
library(sjPlot)
library(sjmisc)
library(lme4)
library(pander)
library(car)
library(olsrr)
library(systemfit)
library(het.test)
panderOptions('knitr.auto.asis', FALSE)

#-- White test function
white.test <- function(lmod,data=d){
  u2 <- lmod$residuals^2
  y <- fitted(lmod)
  Ru2 <- summary(lm(u2 ~ y + I(y^2)))$r.squared
  LM <- nrow(data)*Ru2
  p.value <- 1-pchisq(LM, 2)
  data.frame("Test statistic"=LM,"P value"=p.value)
}

#-- funzione per ottenere osservazioni outlier univariate
FIND_EXTREME_OBSERVARION <- function(x,sd_factor=2){
  which(x>mean(x)+sd_factor*sd(x) | x<mean(x)-sd_factor*sd(x))
}

#-- import dei dati
```

```r
ABSOLUTE_PATH <- "C:\\Users\\sbarberis\\Dropbox\\MODELLI STATISTICI"
d <- read.csv(paste0(ABSOLUTE_PATH,"\\esercizi (3) copia\\3.multilevel\\Exam.txt"),sep=" ")

#-- Fisso la decima scuola come riferimento
d$school <- factor(d$school)
contrasts(d$school) <- contr.treatment(levels(d$school),base=which(levels(d$school) == '65'))


#-- vettore di variabili numeriche presenti nei dati
VAR_NUMERIC <- c("normexam","schavg","standLRT")

#-- print delle prime 6 righe del dataset
pander(head(d))
```

Table 1: Table continues below

id	school	normexam	schgend	schavg	vr	intake	standLRT
1	1	0.2613	mixed	0.1662	mid 50%	bottom 25%	0.6191
2	1	0.1341	mixed	0.1662	mid 50%	mid 50%	0.2058
3	1	-1.724	mixed	0.1662	mid 50%	top 25%	-1.365
4	1	0.9676	mixed	0.1662	mid 50%	mid 50%	0.2058
5	1	0.5443	mixed	0.1662	mid 50%	mid 50%	0.3711
6	1	1.735	mixed	0.1662	mid 50%	bottom 25%	2.189

sex	type	student
F	Mxd	143
F	Mxd	145
M	Mxd	142
F	Mxd	141
F	Mxd	138
M	Mxd	155

STATISTICHE DESCRITTIVE

Si propongono la matrice di correlazione tra le variabili e alcune descrittive di base.

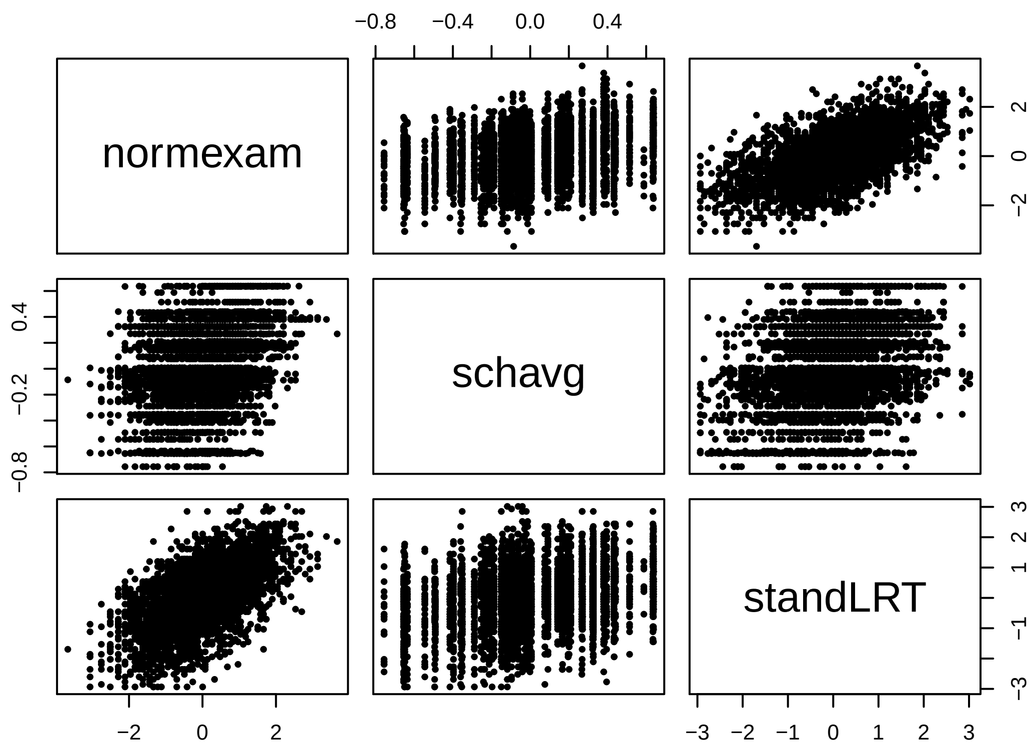
```r
#-- R CODE
pander(summary(d[,VAR_NUMERIC]),big.mark=",") #-- statistiche descrittive
```

normexam	schavg	standLRT
Min. :-3.666072	Min. :-0.75596	Min. :-2.93495
1st Qu.:-0.699505	1st Qu.:-0.14934	1st Qu.:-0.62071
Median : 0.004322	Median :-0.02020	Median : 0.04050
Mean :-0.000114	Mean : 0.00181	Mean : 0.00181
3rd Qu.: 0.678759	3rd Qu.: 0.21053	3rd Qu.: 0.61906
Max. : 3.666091	Max. : 0.63766	Max. : 3.01595

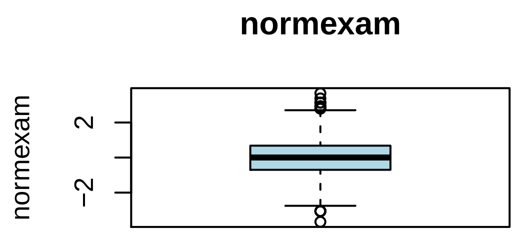
```r
pander(cor(d[,VAR_NUMERIC]),big.mark=",") #-- matrice di correlazione
```

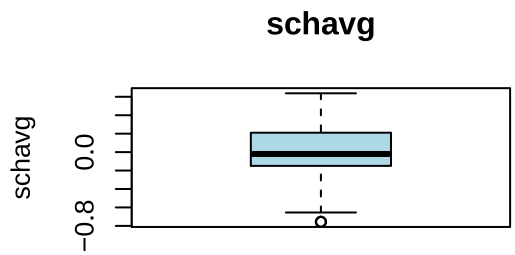
	normexam	schavg	standLRT
normexam	1	0.2879	0.5916
schavg	0.2879	1	0.317
standLRT	0.5916	0.317	1

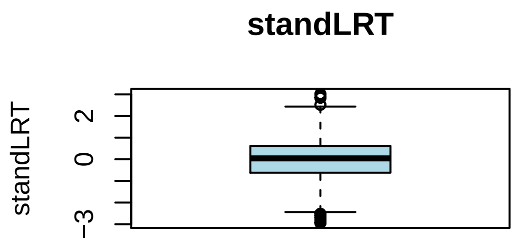
```r
plot(d[,VAR_NUMERIC],pch=19,cex=.5) #-- scatter plot multivariato
```



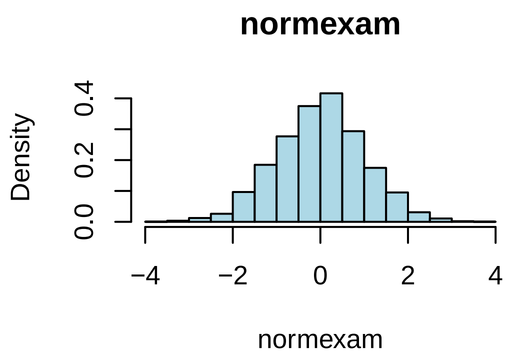

```r
par(mfrow=c(2,2))
for(i in VAR_NUMERIC){
  boxplot(d[,i],main=i,col="lightblue",ylab=i)
}
par(mfrow=c(2,2))
```

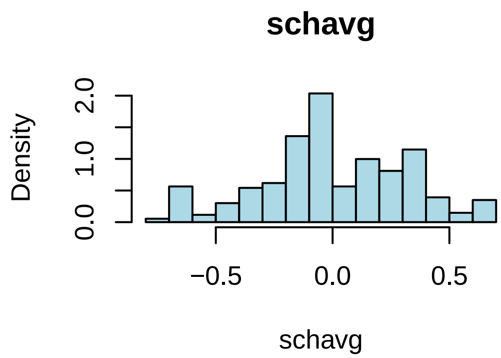


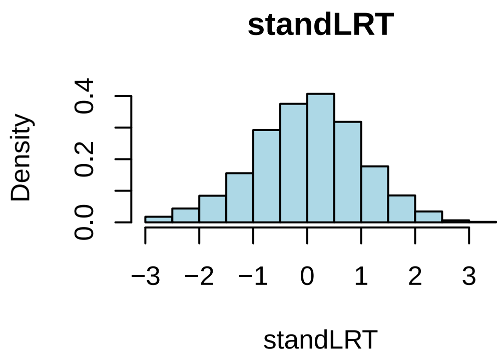





```
for(i in VAR_NUMERIC){
  hist(d[,i],main=i,col="lightblue",xlab=i,freq=F)
}
```

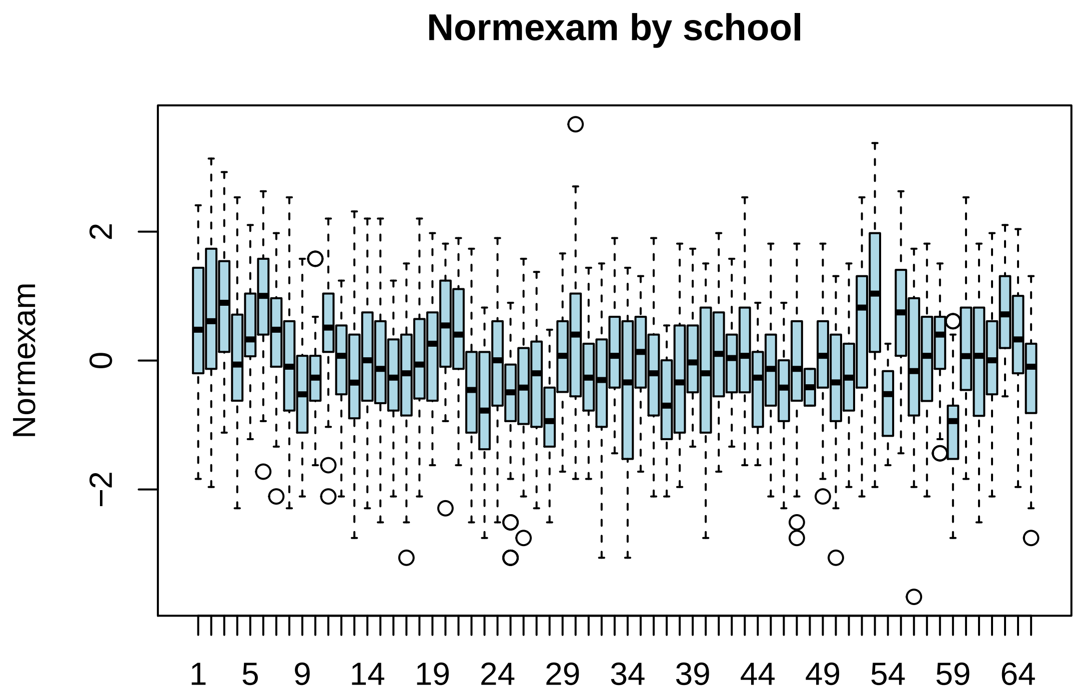






Si propongono poi i box-plot per la variabile dipendente “normexam” per scuola:

```
#-- R CODE
boxplot(d$normexam~d$school,main="Normexam by school",col="lightblue",ylab="Normexam")
```



ANALISI DELLA VARIANZA (EFFETTI FISSI)

Si propone innanziutto un modello di varianza a effetti fissi.

```
#-- R CODE
mod1 <- lm(normexam ~ school,d)
pander(summary(mod1),big.mark=",")
```

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	-0.3087	0.1029	-2.999	0.002728
school1	0.8099	0.149	5.435	5.823e-08
school2	1.092	0.1613	6.77	1.479e-11
school3	1.164	0.164	7.098	1.492e-12
school4	0.3823	0.146	2.618	0.008881
school5	0.7123	0.1866	3.817	0.0001369
school6	1.253	0.1456	8.609	1.046e-17
school7	0.7002	0.1422	4.923	8.874e-07
school8	0.2605	0.1375	1.894	0.05824
school9	-0.127	0.1885	-0.6737	0.5005
school10	0.0393	0.166	0.2367	0.8129
school11	0.8659	0.1558	5.558	2.903e-08
school12	0.2362	0.1692	1.396	0.1628
school13	0.06294	0.1544	0.4076	0.6836

	Estimate	Std. Error	t value	Pr(>|t|)
school14	0.3243	0.122	2.658	0.007883
school15	0.2681	0.1411	1.9	0.05753
school16	0.05457	0.1422	0.3837	0.7012
school17	0.06326	0.1316	0.4806	0.6308
school18	0.3025	0.1329	2.277	0.02287
school19	0.515	0.1613	3.193	0.001418
school20	0.8078	0.1798	4.492	7.246e-06
school21	0.689	0.149	4.623	3.895e-06
school22	-0.1895	0.1415	-1.34	0.1805
school23	-0.4289	0.2022	-2.122	0.03392
school24	0.3213	0.1831	1.755	0.07927
school25	-0.3044	0.149	-2.043	0.04115
school26	-0.08333	0.148	-0.5631	0.5734
school27	-0.02021	0.1798	-0.1124	0.9105
school28	-0.564	0.1596	-3.534	0.0004137
school29	0.3722	0.146	2.549	0.01085
school30	0.6441	0.1754	3.671	0.0002443
school31	0.07201	0.167	0.4311	0.6664
school32	-0.06264	0.1754	-0.3571	0.7211
school33	0.3858	0.147	2.625	0.008707
school34	-0.06221	0.2079	-0.2993	0.7647
school35	0.3926	0.1814	2.165	0.03048
school36	0.05948	0.1507	0.3947	0.6931
school37	-0.3567	0.2217	-1.609	0.1077
school38	0.02114	0.1622	0.1304	0.8963
school39	0.347	0.1681	2.064	0.03905
school40	0.04901	0.1501	0.3265	0.7441
school41	0.3816	0.1572	2.427	0.01527
school42	0.3314	0.1588	2.087	0.03696
school43	0.4563	0.1565	2.915	0.003572
school44	-0.125	0.1996	-0.6265	0.531
school45	0.08318	0.1631	0.5101	0.61
school46	-0.1487	0.1443	-1.031	0.3026
school47	0.1864	0.1447	1.288	0.1977
school48	-0.1056	0.6591	-0.1602	0.8727
school49	0.3556	0.1345	2.643	0.008247
school50	-0.01309	0.149	-0.08786	0.93
school51	0.03062	0.1588	0.1928	0.8471
school52	0.8421	0.1565	5.38	7.861e-08
school53	1.312	0.1507	8.708	4.445e-18
school54	-0.3249	0.3414	-0.9518	0.3413
school55	1.026	0.165	6.218	5.562e-10
school56	0.2744	0.1814	1.513	0.1304
school57	0.3098	0.1551	1.997	0.04585
school58	0.5906	0.1831	3.226	0.001264
school59	-0.7404	0.1692	-4.376	1.242e-05
school60	0.5041	0.1456	3.463	0.0005402
school61	0.2564	0.1544	1.66	0.09695
school62	0.3467	0.1501	2.309	0.02098
school63	1.044	0.1971	5.298	1.232e-07
school64	0.6525	0.158	4.13	3.709e-05

Table 6: Fitting linear model: normexam ~ school

Observations	Residual Std. Error	R^2	Adjusted R^2
4059	0.9207	0.1639	0.1505

```
pander(anova(mod1),big.mark=",")
```

Table 7: Analysis of Variance Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
school	64	663.6	10.37	12.23	9.336e-112
Residuals	3,994	3,386	0.8477	NA	NA

Il test F ci mostra che esiste una struttura gerarchica dei dati in quanto è respinta l'ipotesi nulla che il modello non interpreti i dati e che le scuole non siano significative nello spiegare i risultati scolastici. Si possono quindi presentare i valori delle intercette relative alle scuole che sono calcolate come differenza dai valori attesi generali per il modello e quindi possono essere positive per le scuole più efficaci che la media delle scuole e negativi per quelle meno efficaci. Si può inoltre costruire una graduatoria dell'efficacia delle singole scuole. A questo punto si propone l'empty model.

REGRESSIONE MULTILEVEL: Empty Model

```
#-- R CODE
mod1 <- lmer(normexam ~ (1| school),d,REML=T)
summary(mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: normexam ~ (1 | school)
##    Data: d
##
## REML criterion at convergence: 11014.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.9471 -0.6491  0.0117  0.6987  3.6572
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  school   (Intercept) 0.1716   0.4142
##  Residual             0.8478   0.9207
## Number of obs: 4059, groups:  school, 65
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept) -0.01325    0.05405  -0.245
```

```
pander(Anova(mod1, type="III"),big.mark=",")
```

```
##
## --------------------------------------------
```

```
##              Chisq    Df   Pr(>Chisq)
## ----------------- --------- ---- ------------
##  **(Intercept)**   0.0601    1      0.8063
## ---------------------------------------------
##
## Table: Analysis of Deviance Table (Type III Wald chisquare tests)
```

```
mod1_null <- lm(normexam ~ 1,d)
pander(anova(mod1,mod1_null),big.mark=",")
```

```
##
## ---------------------------------------------------------------------------
##             Df    AIC      BIC      logLik    deviance   Chisq   Chi Df
## --------------- ---- --------- --------- --------- ----------- ------- --------
##  **mod1_null**   2    11,513   11,526   -5,755    11,509     NA       NA
##
##    **mod1**      3    11,017   11,036   -5,505    11,011    498.7     1
## ---------------------------------------------------------------------------
##
## Table: Data: d (continued below)
##
##
## ----------------------------
##             Pr(>Chisq)
## --------------- ------------
##  **mod1_null**       NA
##
##    **mod1**      1.808e-110
## ----------------------------
```

```
pander(data.frame("ICC"=icc(mod1)),big.mark=",") #-- ICC
```

```
##
## ---------------------
##           ICC
## ------------ --------
##  **school**   0.1683
## ---------------------
```

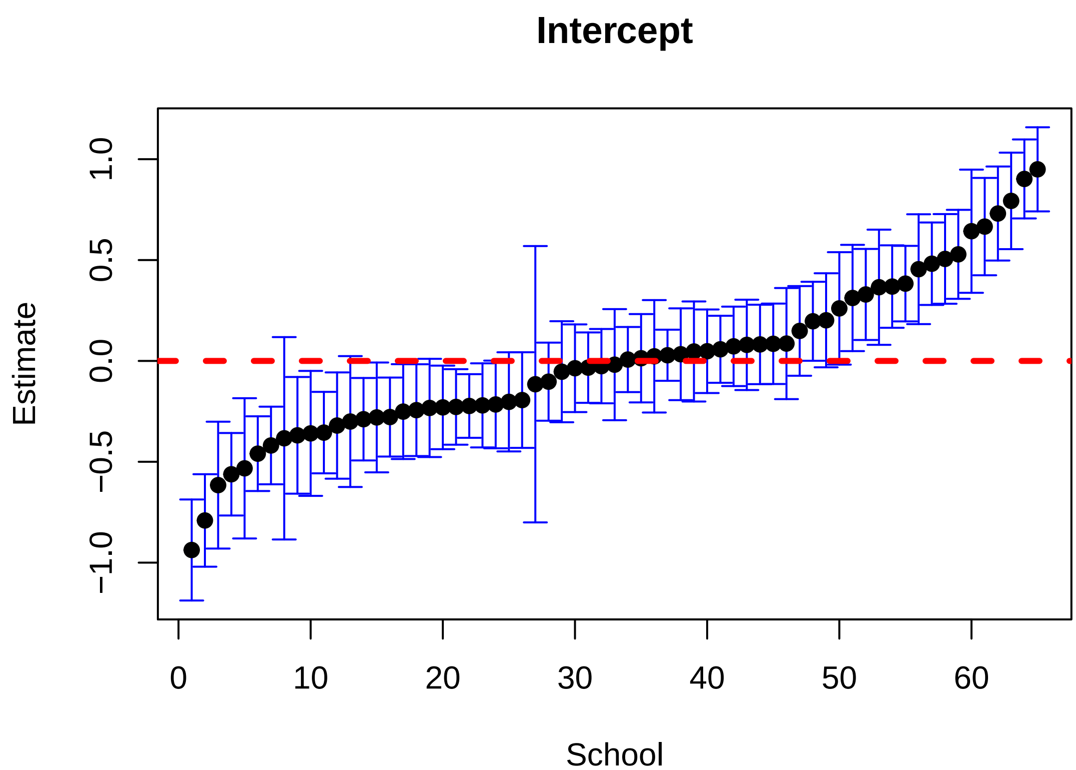
```
res <- sjp.lmer(mod1, type = "re.qq", sort.est = "sort.all",show.values=T,title="T",prnt.plot=F)
res$data$lower <- res$data$y-res$data$ci
res$data$upper <- res$data$y+res$data$ci

pander(res$data[1:10,c("ID","y","upper","lower")])
```

```
##
## -----------------------------------------------------
##                 ID      y        upper      lower
## ------------------- ---- --------- ---------- ---------
##  **(Intercept)59**   59   -0.9373   -0.6869    -1.188
##
##  **(Intercept)28**   28   -0.7909   -0.5616     -1.02
##
##  **(Intercept)23**   23   -0.6157   -0.3013    -0.9302
##
##  **(Intercept)25**   25   -0.5618   -0.3574    -0.7662
```

```
##
##   **(Intercept)37**   37   -0.5325   -0.1848    -0.8802
##
##   **(Intercept)22**   22   -0.4597   -0.2745    -0.6449
##
##   **(Intercept)46**   46   -0.4192   -0.2268    -0.6116
##
##   **(Intercept)54**   54   -0.3835    0.1181    -0.8852
##
##   **(Intercept)9**    9    -0.3688   -0.07964   -0.658
##
##   **(Intercept)44**   44   -0.3593   -0.04951   -0.669
## -----------------------------------------------------
```

```r
plotCI(1:nrow(res$data),res$data$y,ui=res$data$upper, li=res$data$lower,pch=19,scol="blue",xlab="School
abline(h=mean(res$data$y),col=2,lwd=3,lty=2)
```



Il modello interpreta bene i dati ma l'intercetta, unico effetto fisso non è significativo. Il coefficiente di correlazione intraclasse non è insignifiante benchè non particolaremnte elevato trattandosi di un modello empty.

Si propongono quindi gli effetti casuali relativi ad ogni scuola che, come è si è detto sono espressi in termini di differenza dal valore atteso generale. Alcuni sono positivi altri negativi. Tra essi alcuni sono significativi sia pure per diversi livelli di significatività, vale come è noto il confronto tra diverse scuole in termini di efficacia non è svolto sulla base dei valori attesi ma in termini di intervalli di confidenza che appaiono nelle ultime due colonne: una scuola A è più efficace di un'altra B se l'estremo inferiore dell' intervallo di confidenza di A è superiore all'estremo superiore dell'intervallo di confidenza di B.

Si propone ora un mixed model con variabili esplicative "sex", "intake" e "standLRT".

REGRESSIONE MULTILEVEL: Random Intercept

```
#-- R CODE
mod1 <- lmer(normexam ~ sex + intake + standLRT + (1| school),d,REML=T)
summary(mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: normexam ~ sex + intake + standLRT + (1 | school)
##    Data: d
##
## REML criterion at convergence: 9136.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6037 -0.6347  0.0385  0.6743  3.3614
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  school   (Intercept) 0.08419  0.2901
##  Residual             0.53322  0.7302
## Number of obs: 4059, groups:  school, 65
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)    0.41685    0.04707   8.855
## sexM          -0.16303    0.03192  -5.108
## intakemid 50% -0.41504    0.03178 -13.061
## intaketop 25% -0.76030    0.05359 -14.188
## standLRT       0.38799    0.01674  23.173
##
## Correlation of Fixed Effects:
##              (Intr) sexM   int50% int25%
## sexM         -0.285
## intakemd50% -0.488 -0.012
## intaketp25% -0.409 -0.018  0.660
## standLRT    -0.322  0.032  0.543  0.678
```

```
pander(Anova(mod1, type="III"),big.mark=",")
```

```
##
## -------------------------------------------
##               Chisq   Df   Pr(>Chisq)
## ----------------- ------- ---- ------------
##  **(Intercept)**   78.42   1    8.342e-19
##
##      **sex**       26.09   1    3.262e-07
##
##    **intake**      225.5   2    1.061e-49
##
##   **standLRT**      537    1    8.431e-119
## -------------------------------------------
```

```
##
## Table: Analysis of Deviance Table (Type III Wald chisquare tests)
```

```r
pander(data.frame("ICC"=icc(mod1)),big.mark=",") #-- ICC
```

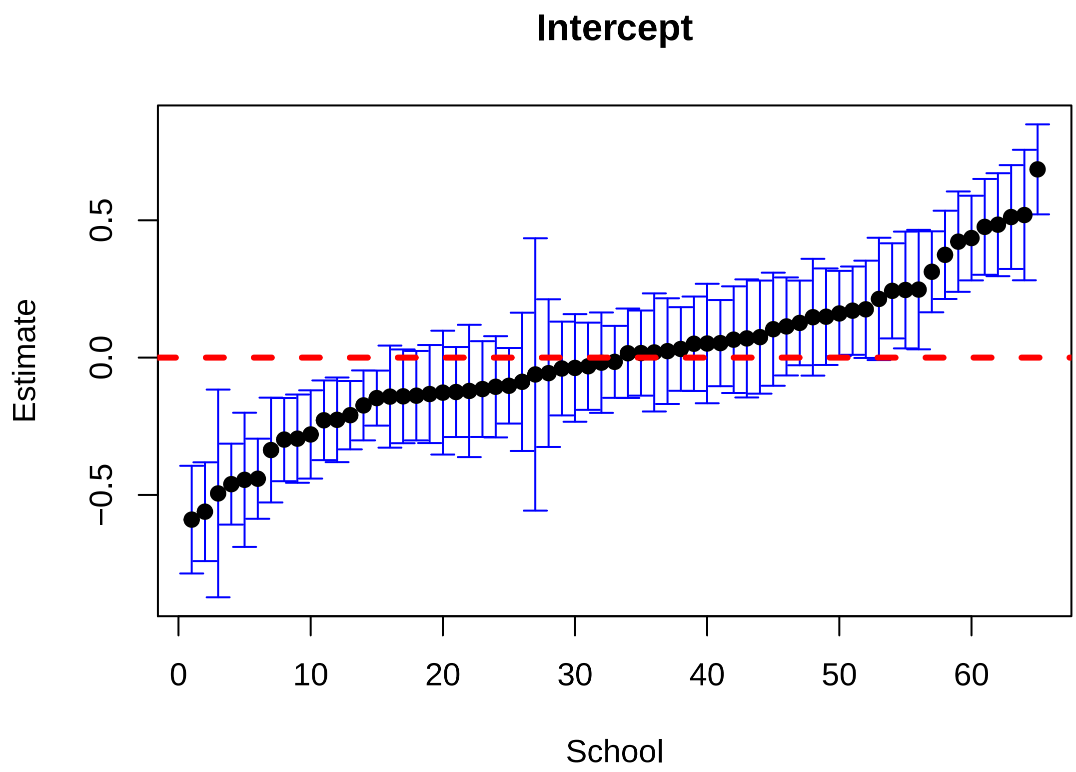
```
##
## ---------------------
##           ICC
## ------------ --------
##  **school**   0.1364
## ---------------------
```

```r
res <- sjp.lmer(mod1, type = "re.qq", sort.est = "sort.all",show.values=T,title="T",prnt.plot=F)
res$data$lower <- res$data$y-res$data$ci
res$data$upper <- res$data$y+res$data$ci

pander(res$data[1:10,c("ID","y","upper","lower")])
```

```
##
## -----------------------------------------------------
##                 ID      y        upper     lower
## ------------------- ---- --------- --------- ---------
##  **(Intercept)59**   59   -0.5898   -0.3938   -0.7858
##
##  **(Intercept)28**   28   -0.5609   -0.381    -0.7407
##
##  **(Intercept)54**   54   -0.4944   -0.1164   -0.8724
##
##  **(Intercept)16**   16   -0.4606   -0.3133   -0.608
##
##  **(Intercept)23**   23   -0.4449   -0.2006   -0.6892
##
##  **(Intercept)22**   22   -0.4409   -0.2951   -0.5867
##
##  **(Intercept)10**   10   -0.3365   -0.1458   -0.5272
##
##  **(Intercept)46**   46   -0.2985   -0.1471   -0.4499
##
##  **(Intercept)25**   25   -0.295    -0.1343   -0.4557
##
##  **(Intercept)50**   50   -0.2797   -0.119    -0.4404
## -----------------------------------------------------
```

```r
plotCI(1:nrow(res$data),res$data$y,ui=res$data$upper, li=res$data$lower,pch=19,scol="blue",xlab="School
abline(h=mean(res$data$y),col=2,lwd=3,lty=2)
```



Tutte i variabili esplicative sono risultano significativi. Si passa ora al mixed model con tutte le variabili esplicative.

```r
#-- R CODE
mod1 <- lmer(normexam ~ vr + intake + sex + type + schgend + (1| school),d,REML=T)
summary(mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: normexam ~ vr + intake + sex + type + schgend + (1 | school)
##    Data: d
##
## REML criterion at convergence: 9627.6
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6134 -0.6406  0.0333  0.6716  3.4815
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  school   (Intercept) 0.07321  0.2706
##  Residual             0.60366  0.7770
## Number of obs: 4059, groups:  school, 65
##
## Fixed effects:
##                  Estimate Std. Error t value
```

```
## (Intercept)    0.49277    0.09331    5.28
## vrmid 50%      0.16717    0.09576    1.75
## vrtop 25%      0.42040    0.10773    3.90
## intakemid 50% -0.81030    0.02843  -28.50
## intaketop 25% -1.59270    0.04201  -37.92
## sexM          -0.18117    0.03527   -5.14
## typeSngl       0.20482    0.10693    1.92
## schgendgirls  -0.09097    0.11985   -0.76
##
## Correlation of Fixed Effects:
##             (Intr) vrm50% vrt25% int50% int25% sexM   typSng
## vrmid 50%   -0.765
## vrtop 25%   -0.675  0.649
## intakemd50% -0.225  0.020  0.053
## intaketp25% -0.192  0.053  0.081  0.474
## sexM        -0.198  0.013  0.018 -0.037 -0.051
## typeSngl    -0.269  0.062  0.063  0.012  0.005 -0.158
## schgendgrls  0.009 -0.053 -0.110 -0.022 -0.011  0.292 -0.728
## fit warnings:
## fixed-effect model matrix is rank deficient so dropping 1 column / coefficient
```

```r
pander(Anova(mod1, type="III"),big.mark=",")
```

```
##
## ---------------------------------------------
##               Chisq    Df   Pr(>Chisq)
## ----------------- -------- ---- ------------
##  **(Intercept)**   27.89    1    1.285e-07
##
##      **vr**        16.29    2    0.0002899
##
##    **intake**      1,581    2        0
##
##      **sex**       26.38    1    2.798e-07
##
##     **type**       3.669    1     0.05543
##
##    **schgend**     0.5761   1     0.4478
## ---------------------------------------------
##
## Table: Analysis of Deviance Table (Type III Wald chisquare tests)
```

```r
pander(data.frame("ICC"=icc(mod1)),big.mark=",") #-- ICC
```

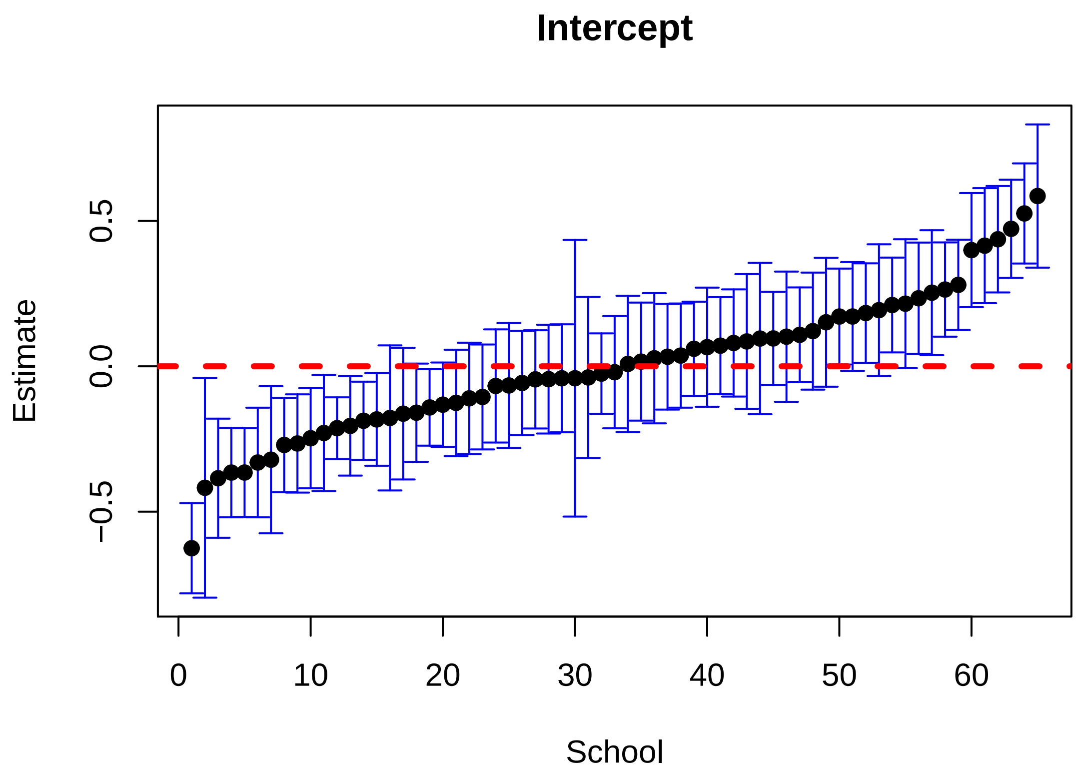
```
##
## ---------------------
##           ICC
## ------------ --------
##  **school**   0.1082
## ---------------------
```

```r
res <- sjp.lmer(mod1, type = "re.qq", sort.est = "sort.all",show.values=T,title="T",prnt.plot=F)
res$data$lower <- res$data$y-res$data$ci
res$data$upper <- res$data$y+res$data$ci

pander(res$data[1:10,c("ID","y","upper","lower")])
```

```
##
## -------------------------------------------------------
##                  ID       y        upper      lower
## ------------------- ---- --------- ---------- ---------
##  **(Intercept)16**   16   -0.6256   -0.4704    -0.7808
##
##  **(Intercept)54**   54   -0.4178   -0.04002   -0.7957
##
##  **(Intercept)59**   59   -0.3849    -0.18     -0.5898
##
##  **(Intercept)22**   22   -0.3655   -0.2119    -0.5191
##
##  **(Intercept)15**   15   -0.3653   -0.2124    -0.5182
##
##  **(Intercept)28**   28   -0.3308   -0.1423    -0.5194
##
##  **(Intercept)23**   23   -0.3212   -0.06825   -0.5741
##
##  **(Intercept)65**   65   -0.2705   -0.1084    -0.4326
##
##  **(Intercept)25**   25   -0.2654   -0.09642   -0.4343
##
##  **(Intercept)36**   36   -0.2474   -0.07526   -0.4196
## -------------------------------------------------------
```

```
plotCI(1:nrow(res$data),res$data$y,ui=res$data$upper, li=res$data$lower,pch=19,scol="blue",xlab="School
abline(h=mean(res$data$y),col=2,lwd=3,lty=2)
```



Il modello interpreta bene i dati e il coefficiente intraclasse diminuisce leggermente rispetto al precedente modello. Si passa ora la modello total effects che contiene due variabili esplicative, una con parametro casuale “standLRT” e l’altra con effetto fisso “schgend”.

REGRESSIONE MULTILEVEL: Random Slope

```
#-- R CODE
mod1 <- lmer(normexam ~ standLRT + schavg + (standLRT| school),d,REML=T)
summary(mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: normexam ~ standLRT + schavg + (standLRT | school)
##    Data: d
##
## REML criterion at convergence: 9323.9
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.8294 -0.6317  0.0326  0.6851  3.4363
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  school   (Intercept) 0.07720  0.2778
##           standLRT    0.01532  0.1238   0.37
##  Residual             0.55360  0.7440
```

```
## Number of obs: 4059, groups:  school, 65
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.001423   0.037255  -0.038
## standLRT     0.552242   0.020353  27.133
## schavg       0.294731   0.107267   2.748
##
## Correlation of Fixed Effects:
##          (Intr) stnLRT
## standLRT  0.266
## schavg    0.089 -0.085
```

```
pander(Anova(mod1, type="III"),big.mark=",")
```

```
##
## ---------------------------------------------
##                Chisq    Df   Pr(>Chisq)
## ----------------- --------- ---- ------------
##  **(Intercept)**   0.00146   1      0.9695
##
##   **standLRT**      736.2    1    4.037e-162
##
##    **schavg**       7.55     1     0.006002
## ---------------------------------------------
##
## Table: Analysis of Deviance Table (Type III Wald chisquare tests)
```

```
pander(data.frame("ICC"=icc(mod1)),big.mark=",") #-- ICC
```

```
##
## ---------------------
##           ICC
## ------------ --------
##  **school**   0.1224
## ---------------------
```

```
res <- sjp.lmer(mod1, type = "re.qq", sort.est = "sort.all",show.values=T,title="T",prnt.plot=F)
res$data$lower <- res$data$y-res$data$ci
res$data$upper <- res$data$y+res$data$ci

res_int <- subset(res$data,ind=="(Intercept)")
res_hw <- subset(res$data,ind=="standLRT")

pander(res_int[1:10,c("ID","y","upper","lower")])
```

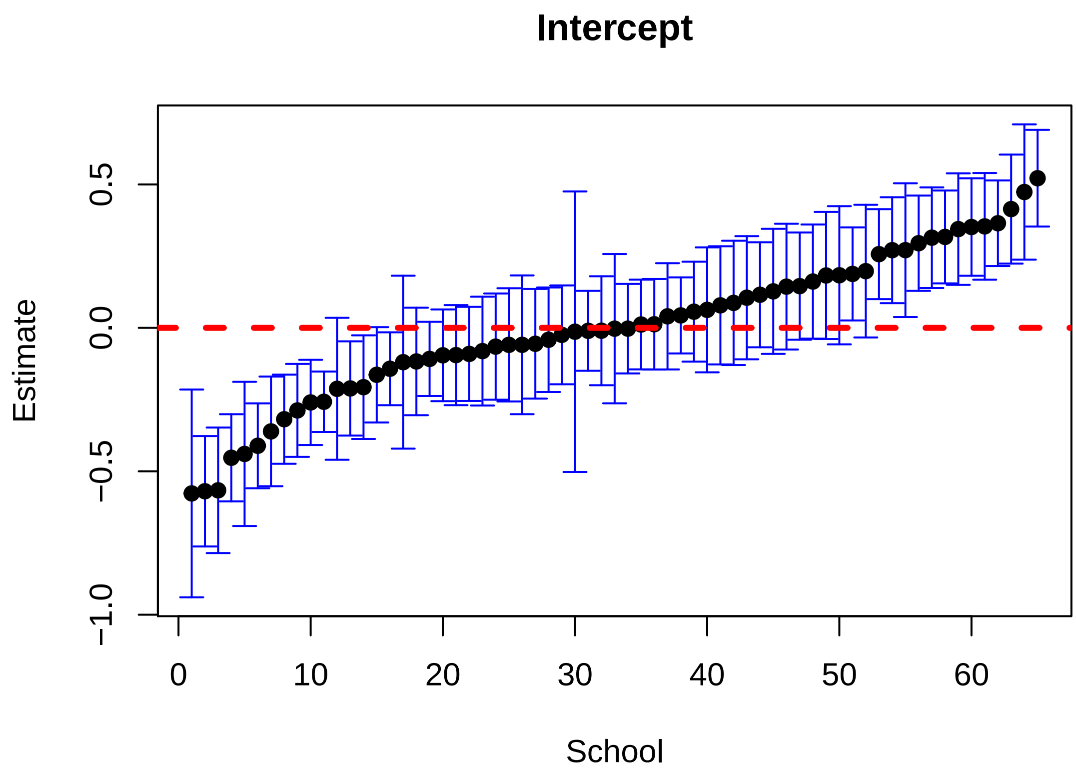
```
##
## ------------------------------------------------------
##                 ID      y       upper     lower
## ------------------- ---- --------- --------- ---------
##  **(Intercept)54**   54   -0.5771   -0.2151   -0.9392
##
##  **(Intercept)28**   28   -0.5696   -0.3772   -0.7619
##
##  **(Intercept)59**   59   -0.5663   -0.3475   -0.7852
##
```

```
## **(Intercept)16**   16   -0.4529   -0.301    -0.6048
##
## **(Intercept)23**   23   -0.4396   -0.1882   -0.691
##
## **(Intercept)22**   22   -0.4111   -0.2632   -0.5591
##
## **(Intercept)10**   10   -0.361    -0.1699   -0.5521
##
## **(Intercept)46**   46   -0.3186   -0.1631   -0.474
##
## **(Intercept)50**   50   -0.2878   -0.1257   -0.4499
##
## **(Intercept)15**   15    -0.26    -0.1115   -0.4085
## ---------------------------------------------------------
```

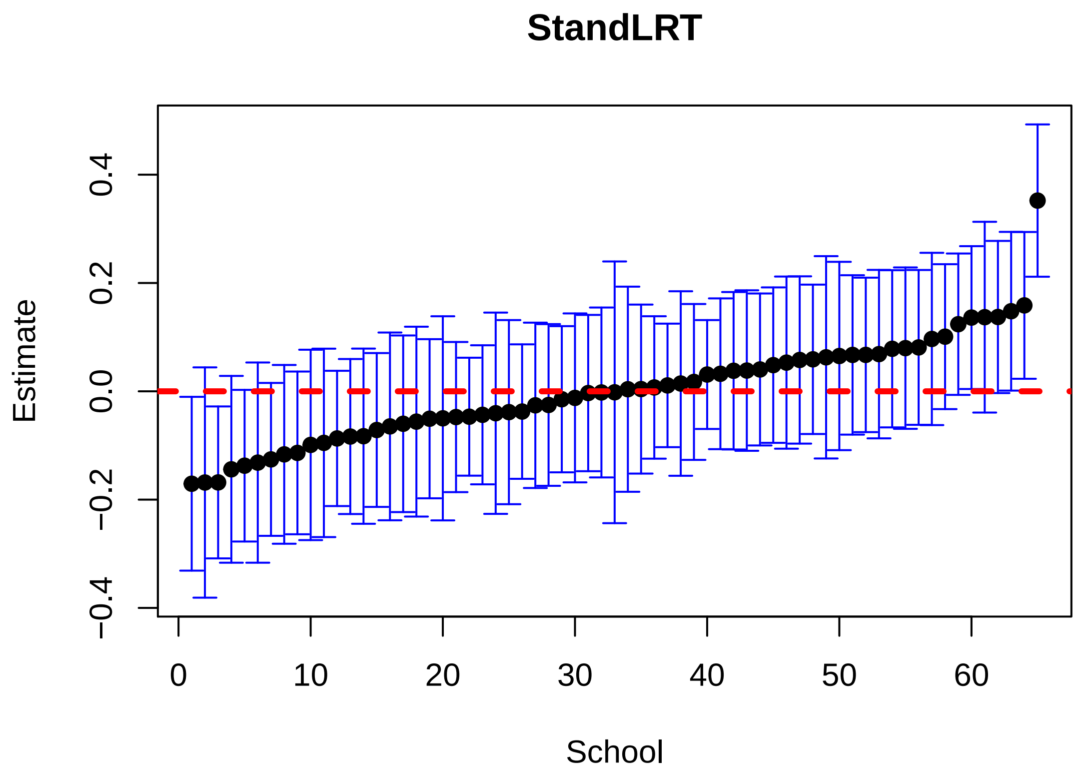
```
pander(res_hw[1:10,c("ID","y","upper","lower")])
```

```
##
## --------------------------------------------------------
##              ID      y         upper      lower
## ---------------- ---- ---------- ---------- ---------
## **standLRT28**   28   -0.1707    -0.01023   -0.3311
##
## **standLRT54**   54   -0.1684    0.04417    -0.381
##
## **standLRT7**    7    -0.1682    -0.02794   -0.3085
##
## **standLRT10**   10   -0.1441    0.02842    -0.3165
##
## **standLRT51**   51   -0.1374    0.002659   -0.2774
##
## **standLRT37**   37   -0.1317    0.05319    -0.3165
##
## **standLRT16**   16   -0.1257    0.01543    -0.2669
##
## **standLRT59**   59   -0.1164    0.04859    -0.2815
##
## **standLRT18**   18   -0.1137     0.0365    -0.264
##
## **standLRT23**   23   -0.09898   0.07682    -0.2748
## --------------------------------------------------------
```

```
plotCI(1:nrow(res_int),res_int$y,ui=res_int$upper, li=res_int$lower,pch=19,scol="blue",xlab="School",yla
abline(h=mean(res_int$y),col=2,lwd=3,lty=2)
```




```
plotCI(1:nrow(res_hw),res_hw$y,ui=res_hw$upper, li=res_hw$lower,pch=19,scol="blue",xlab="School",ylab="E
abline(h=mean(res_hw$y),col=2,lwd=3,lty=2)
```



Il modello interpreta bene i dati e sia i parametri casuali relativi a intercetta che la variabile esplicativa risultano significativi come anche il coefficiente di correlazione di valore positivo. I parametri fissi (la parte fissa del parametro casuale relativo a "standLRT" scomponibile in una parte propriamente casuale e una fissa e il parametro relativo a "schavg") sono entrambi significativi. Il test di tipo 3 sugli effetti fissi vine effettuato con la variabile casuale F invece che con la t ma dà risultati identici perchè i valori di F non sono altro che i quadrati dei valori di t.

Si propone ora un altro random model con "intake" come variabile esplicativa con parametro fisso e "standRLT" con parametro casuale.

```
#-- R CODE
mod1 <- lmer(normexam ~ intake + (standLRT| school),d,REML=T)
summary(mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: normexam ~ intake + (standLRT | school)
##    Data: d
##
## REML criterion at convergence: 9221.7
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.7095 -0.6292  0.0366  0.6690  3.5717
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
```

```
##  school   (Intercept) 0.1306   0.3613
##           standLRT    0.1463   0.3824   0.65
##  Residual             0.5232   0.7233
## Number of obs: 4059, groups:  school, 65
##
## Fixed effects:
##               Estimate Std. Error t value
## (Intercept)    0.16073    0.04778   3.364
## intakemid 50% -0.45040    0.03160 -14.252
## intaketop 25% -0.86633    0.05238 -16.538
##
## Correlation of Fixed Effects:
##             (Intr) int50%
## intakemd50% -0.591
## intaketp25% -0.523  0.635
```

```
pander(Anova(mod1, type="III"),big.mark=",")
```

```
##
## -------------------------------------------
##               Chisq   Df   Pr(>Chisq)
## ----------------- -------- ---- ------------
##  **(Intercept)**   11.32    1    0.0007685
##
##    **intake**      297.1    2    3.058e-65
## -------------------------------------------
##
## Table: Analysis of Deviance Table (Type III Wald chisquare tests)
```

```
pander(data.frame("ICC"=icc(mod1)),big.mark=",") #-- ICC
```

```
##
## ---------------------
##           ICC
## ------------ --------
##  **school**   0.1997
## ---------------------
```

```
res <- sjp.lmer(mod1, type = "re.qq", sort.est = "sort.all",show.values=T,title="T",prnt.plot=F)
res$data$lower <- res$data$y-res$data$ci
res$data$upper <- res$data$y+res$data$ci

res_int <- subset(res$data,ind=="(Intercept)")
res_hw <- subset(res$data,ind=="standLRT")

pander(res_int[1:10,c("ID","y","upper","lower")])
```

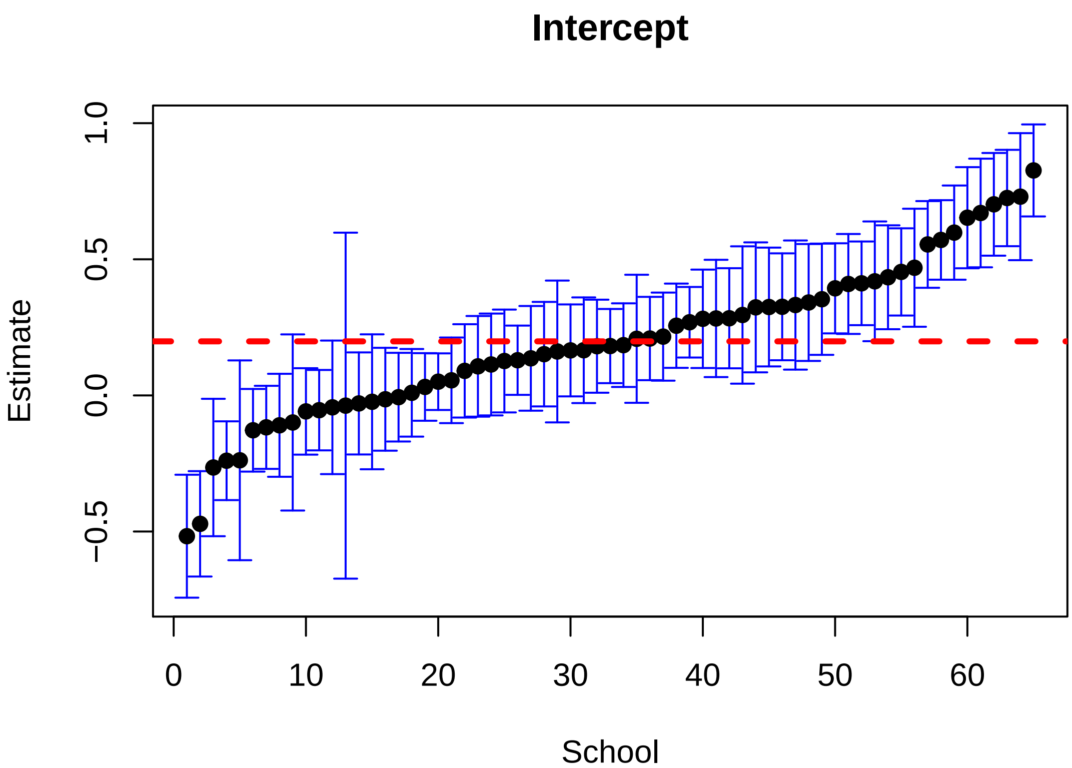
```
##
## ---------------------------------------------------------
##                 ID      y         upper      lower
## ------------------- ---- ---------- ---------- ---------
##  **(Intercept)59**   59   -0.5171    -0.2913    -0.7429
##
##  **(Intercept)28**   28   -0.4718    -0.2782    -0.6653
##
##  **(Intercept)23**   23   -0.2648    -0.01252   -0.5172
```

```
##
##   **(Intercept)22**   22   -0.2399    -0.09522   -0.3846
##
##   **(Intercept)54**   54   -0.2384     0.1286    -0.6053
##
##   **(Intercept)16**   16   -0.1281     0.0237    -0.2799
##
##   **(Intercept)46**   46   -0.1173    0.03524    -0.2699
##
##   **(Intercept)10**   10   -0.1098    0.07932    -0.299
##
##   **(Intercept)37**   37   -0.09942    0.2241    -0.423
##
##   **(Intercept)50**   50   -0.05879   0.09997    -0.2175
## ---------------------------------------------------------
```

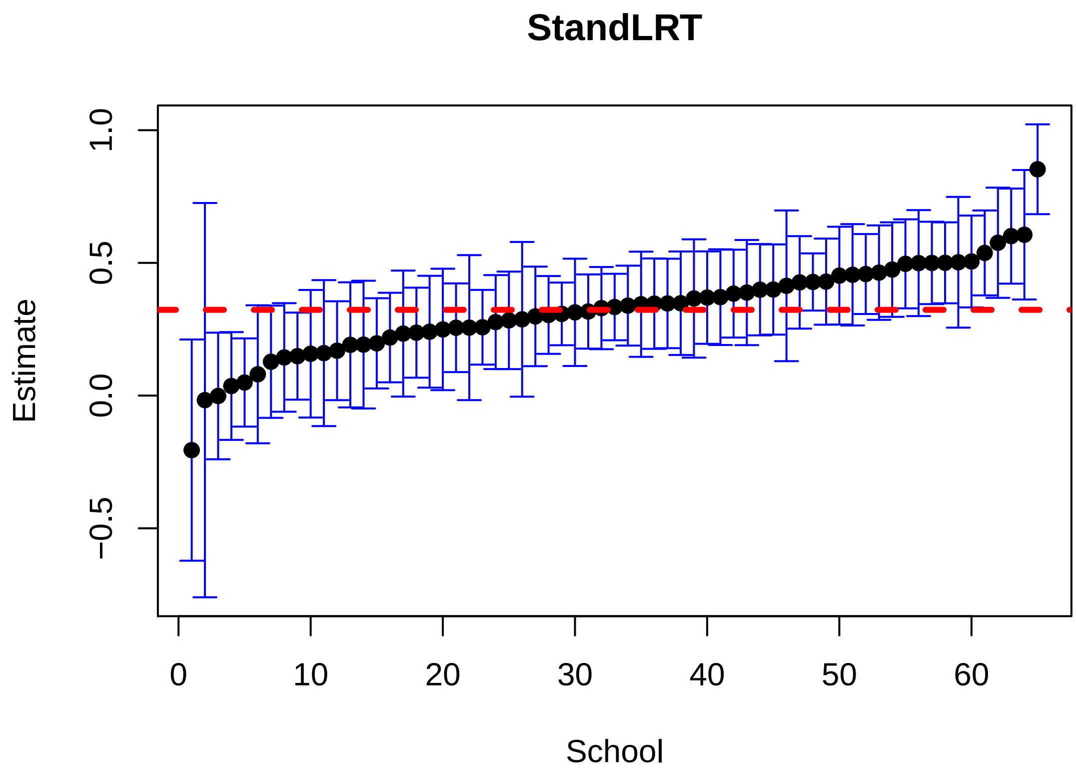
```
pander(res_hw[1:10,c("ID","y","upper","lower")])
```

```
##
## ------------------------------------------------------
##                ID        y        upper     lower
## ---------------- ---- ----------- -------- ----------
##   **standLRT54**   54    -0.2055    0.2113   -0.6222
##
##   **standLRT48**   48   -0.01711    0.7257   -0.7599
##
##   **standLRT10**   10   -0.001452   0.237    -0.2399
##
##   **standLRT28**   28    0.03609    0.2391   -0.1669
##
##   **standLRT7**    7     0.04925    0.2155    -0.117
##
##   **standLRT37**   37    0.08032    0.3403   -0.1796
##
##   **standLRT59**   59    0.1275     0.3389   -0.08394
##
##   **standLRT9**    9     0.1437     0.3483   -0.06081
##
##   **standLRT51**   51    0.1486     0.3126   -0.01529
##
##   **standLRT23**   23    0.1577     0.398    -0.08266
## ------------------------------------------------------
```

```
plotCI(1:nrow(res_int),res_int$y,ui=res_int$upper, li=res_int$lower,pch=19,scol="blue",xlab="School",yla
abline(h=mean(res_int$y),col=2,lwd=3,lty=2)
```




```
plotCI(1:nrow(res_hw),res_hw$y,ui=res_hw$upper, li=res_hw$lower,pch=19,scol="blue",xlab="School",ylab="I
abline(h=mean(res_hw$y),col=2,lwd=3,lty=2)
```



Si propongono ora un ultimo modello:

```r
#-- R CODE
mod1 <- lmer(normexam ~ standLRT + (intake| school),d,REML=T)
summary(mod1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: normexam ~ standLRT + (intake | school)
##    Data: d
##
## REML criterion at convergence: 9174.1
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.6431 -0.6347  0.0358  0.6723  3.5331
##
## Random effects:
##  Groups   Name          Variance Std.Dev. Corr
##  school   (Intercept)   0.3356   0.5793
##           intakemid 50% 0.1720   0.4147   -0.88
##           intaketop 25% 0.6209   0.7880   -0.84  1.00
##  Residual               0.5224   0.7228
## Number of obs: 4059, groups:  school, 65
##
## Fixed effects:
```

```
##             Estimate Std. Error t value
## (Intercept) -0.11490    0.03937  -2.918
## standLRT     0.42511    0.01603  26.516
##
## Correlation of Fixed Effects:
##          (Intr)
## standLRT 0.134
```

```r
pander(Anova(mod1, type="III"),big.mark=",")
```

```
##
## -------------------------------------------
##               Chisq   Df   Pr(>Chisq)
## ----------------- -------- ---- ------------
##  **(Intercept)**   8.517   1     0.003518
##
##   **standLRT**     703.1   1    6.302e-155
## -------------------------------------------
##
## Table: Analysis of Deviance Table (Type III Wald chisquare tests)
```

```r
pander(data.frame("ICC"=icc(mod1)),big.mark=",") #-- ICC
```

```
##
## ---------------------
##           ICC
## ------------ --------
##  **school**   0.3912
## ---------------------
```

```r
res <- sjp.lmer(mod1, type = "re.qq", sort.est = "sort.all",show.values=T,title="T",prnt.plot=F)
res$data$lower <- res$data$y-res$data$ci
res$data$upper <- res$data$y+res$data$ci

res_int <- subset(res$data,ind=="(Intercept)")
res_hw <- subset(res$data,ind=="standLRT")

pander(res_int[1:10,c("ID","y","upper","lower")])
```

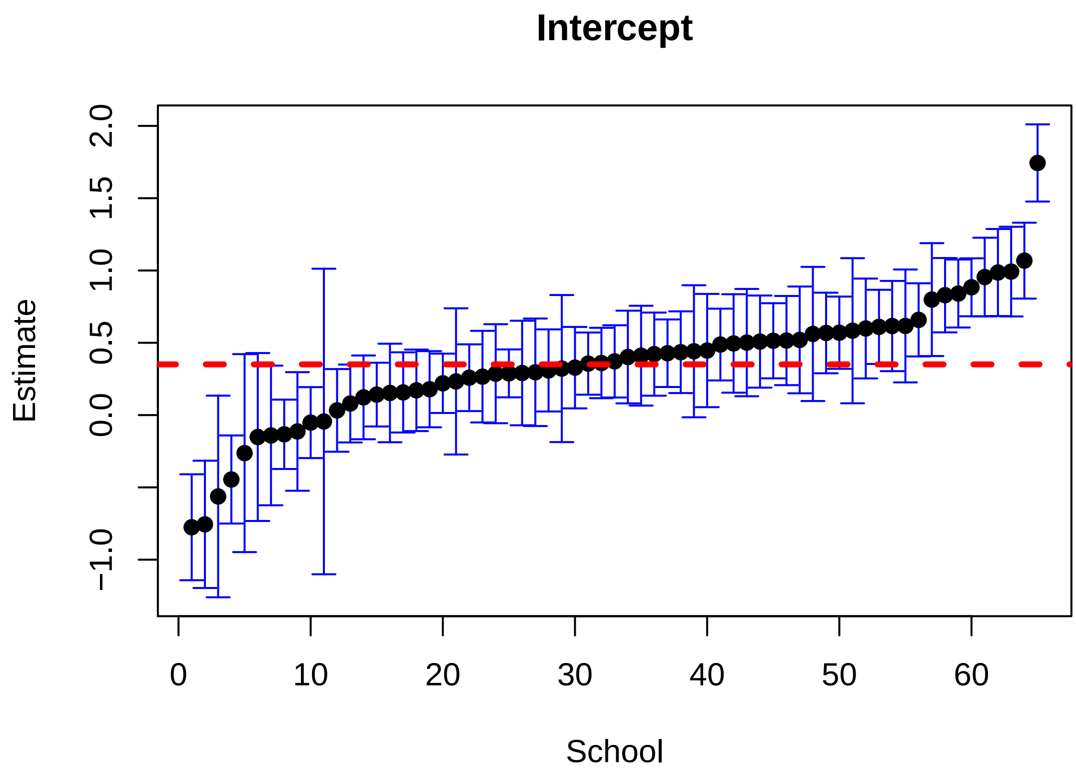
```
##
## ------------------------------------------------------
##                 ID      y        upper     lower
## ------------------- ---- ---------- --------- ---------
##  **(Intercept)28**   28   -0.7758    -0.4094   -1.142
##
##  **(Intercept)59**   59   -0.7556    -0.3154   -1.196
##
##  **(Intercept)54**   54   -0.5628    0.1344     -1.26
##
##  **(Intercept)10**   10   -0.4456    -0.141    -0.7502
##
##  **(Intercept)37**   37   -0.2632    0.4214    -0.9477
##
##  **(Intercept)44**   44   -0.1517     0.429    -0.7324
##
##  **(Intercept)23**   23    -0.141    0.3419    -0.6239
```

```
##
##  **(Intercept)22**   22   -0.1331    0.1068    -0.373
##
##  **(Intercept)9**    9    -0.1133    0.2968    -0.5234
##
##  **(Intercept)16**   16   -0.05207   0.1933    -0.2974
## -------------------------------------------------------
```

```r
pander(res_hw[1:10,c("ID","y","upper","lower")])
```

```
##
## -----------------------------------
##        ID    y     upper   lower
## ---------- ---- ---- ------- -------
##   **NA**    NA   NA    NA      NA
##
##  **NA.1**   NA   NA    NA      NA
##
##  **NA.2**   NA   NA    NA      NA
##
##  **NA.3**   NA   NA    NA      NA
##
##  **NA.4**   NA   NA    NA      NA
##
##  **NA.5**   NA   NA    NA      NA
##
##  **NA.6**   NA   NA    NA      NA
##
##  **NA.7**   NA   NA    NA      NA
##
##  **NA.8**   NA   NA    NA      NA
##
##  **NA.9**   NA   NA    NA      NA
## -----------------------------------
```

```r
plotCI(1:nrow(res_int),res_int$y,ui=res_int$upper, li=res_int$lower,pch=19,scol="blue",xlab="School",yla
abline(h=mean(res_int$y),col=2,lwd=3,lty=2)
```


Intercept
Estimate
School
2.0
1.5
1.0
0.5
0.0
−1.0
0
10
20
30
40
50
60