

## Zkouška z MAS02, 10. 6. 2021

**Popis situace:** 98 náhodně vybraných žáků 3. ročníku základní školy bylo podrobena několika psychologickým testům. K dispozici jsou výsledky těchto pěti testů: vědomosti (ved), podobnosti (pod), počty (poc), slovník (slo), opakování čísel (opak). V každém z těchto testů je možno získat až 20 bodů. Kromě toho byla u každého žáka zhodnocena úroveň jeho hlasitého čtení (varianta 1 – zvládá lehce, varianta 2 – jenom zvládá) a úroveň jeho matematických schopností: varianta 1 – dobrá, varianta 2 – průměrná, varianta 3 - slabá. Zjištěné údaje včetně celkového počtu bodů jsou zaznamenány v datovém souboru `treti_trida.xls`.

### Cíl výzkumu:

1. Zjistit, zda ve výsledcích uvedených pěti psychologických testů existují statisticky významné rozdíly mezi žáky, kteří čtení zvládají lehce a těmi, kteří čtení jenom zvládají.
2. Zjistit, zda na celkový počet bodů z uvedených pěti psychologických testů má vliv úroveň matematických schopností.

(Při testování hypotéz vždy volíme hladinu významnosti 0,05.)

### Výstupy ze statistického software k cíli 1

```
> (m1<-colMeans(treti_trida[treti_trida$cteni=="zvlada_lehce",3:7]))
  ved      pod      poc      slo      opak
10.81818 11.11364 11.18182 10.68182 11.02273
> (m2<-colMeans(treti_trida[treti_trida$cteni=="jenom_zvlada",3:7]))
  ved      pod      poc      slo      opak
9.592593 9.722222 9.407407 9.537037 9.518519

> (var1<-var(treti_trida[treti_trida$cteni=="zvlada_lehce",3:7]))
  ved      pod      poc      slo      opak
ved  4.942918 2.997886 2.824524 3.429175 2.562368
pod  2.997886 8.010042 3.467230 4.641649 3.253171
poc  2.824524 3.467230 7.501057 1.966173 2.623679
slo  3.429175 4.641649 1.966173 6.501057 1.588795
opak 2.562368 3.253171 2.623679 1.588795 6.813425
> (var2<-var(treti_trida[treti_trida$cteni=="jenom_zvlada",3:7]))
  ved      pod      poc      slo      opak
ved  7.642208 4.639413 3.791754 4.449336 2.781272
pod  4.639413 8.506289 3.549266 4.850105 4.335430
poc  3.791754 3.549266 6.736548 2.946890 3.954577
slo  4.449336 4.850105 2.946890 6.630678 1.659679
opak 2.781272 4.335430 3.954577 1.659679 8.593990

> (sqrt(diag(var1))/m1)
  ved      pod      poc      slo      opak
0.2055121 0.2546603 0.2449338 0.2386969 0.2368065
> (sqrt(diag(var2))/m2)
  ved      pod      poc      slo      opak
0.2881864 0.2999884 0.2758981 0.2700011 0.3079839
```

```

> library(mvnTest)
> HZ.test(treti_trida[treti_trida$cteni=="zvlada_lehce",3:7])
      Henze-Zirkler test for Multivariate Normality

data : treti_trida[treti_trida$cteni == "zvlada_lehce", 3:7]

HZ          : 0.7590087
p-value     : 0.800589

Result  : Data are multivariate normal (sig.level = 0.05)

```

```

> HZ.test(treti_trida[treti_trida$cteni=="jenom_zvlada",3:7])
      Henze-Zirkler test for Multivariate Normality

data : treti_trida[treti_trida$cteni == "jenom_zvlada", 3:7]

HZ          : 0.7959112
p-value     : 0.69404

Result  : Data are multivariate normal (sig.level = 0.05)

```

```

> library(biotools)
> boxM(treti_trida[,3:7],grouping=treti_trida$cteni)

      Box's M-test for Homogeneity of Covariance Matrices

```

```

data: treti_trida[, 3:7]
Chi-Sq (approx.) = 9.5393, df = 15, p-value = 0.8477

```

```

> library(ICSNP)
> HotellingsT2(treti_trida[treti_trida$cteni=="zvlada_lehce",3:7],
treti_trida[treti_trida$cteni=="jenom_zvlada",3:7])

```

Hotelling's two sample T2-test

```

data: treti_trida[treti_trida$cteni == "zvlada_lehce", 3:7] and
treti_trida[treti_trida$cteni == "jenom_zvlada", 3:7]
T.2 = 2.579, df1 = 5, df2 = 92, p-value = 0.03138
alternative hypothesis: true location difference is not equal to
c(0,0,0,0,0)

```

```

> n1<-table(treti_trida$cteni)[1]
> n2<-table(treti_trida$cteni)[2]
> n<-n1+n2
> k<-5
> m1<-colMeans(treti_trida[treti_trida$cteni=="zvlada_lehce",3:7])
> m2<-colMeans(treti_trida[treti_trida$cteni=="jenom_zvlada",3:7])
> var1 <- diag(treti_trida[treti_trida$cteni=="zvlada_lehce",3:7])
> var2 <- diag(cov(treti_trida[treti_trida$cteni=="jenom_zvlada",3:7]))
> var <- ( (n1-1)*var1 + (n2-1)*var2 )/(n-2)
> F.stat <- n1*n2*(n-k-1) * (m1-m2)^2 /(var*n*k*(n-2))
> p.hodnota <- 1-pf(F.stat, k, n-k-1)
> kvantil <- qf(0.95, k, n-k-1)
> tab <- round(rbind(F.stat,p.hodnota, kvantil),digits=4)
> rownames(tab) <- c("F","p-hodnota", "kvantil")
> tab

```

	ved	pod	poc	slo	opak
F	1.0850	1.0860	2.0668	0.9266	1.3486
p-hodnota	0.3740	0.3734	0.0766	0.4676	0.2511
kvantil	2.3134	2.3134	2.3134	2.3134	2.3134

## Výstupy ze statistického software k cíli 2

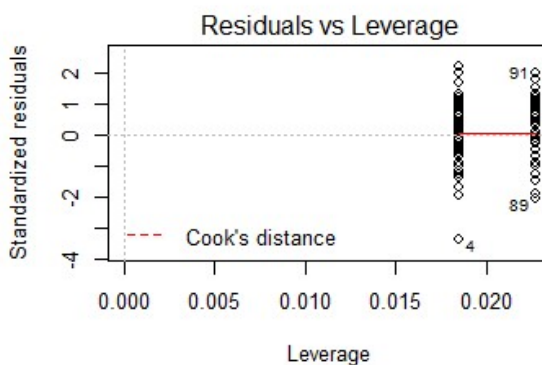
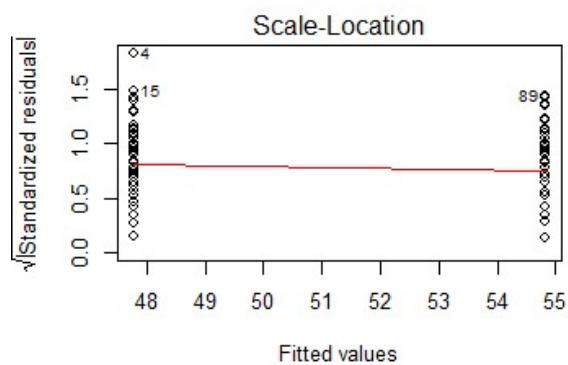
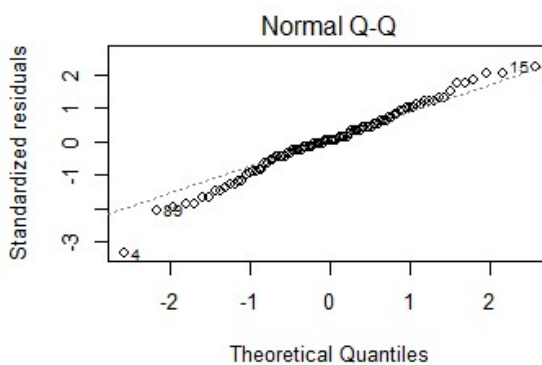
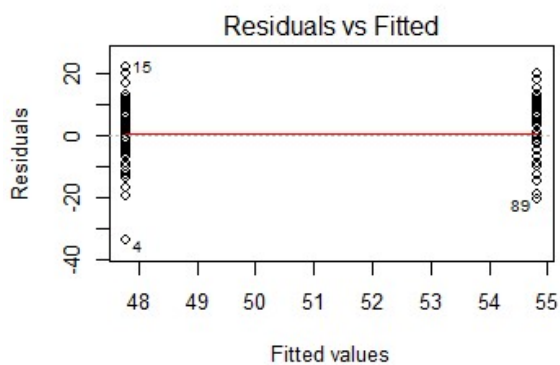
```
> tapply(treti_trida$body,treti_trida$matem,mean)
      dobra prumerna      slaba
54.78846 48.23684 38.75000
> tapply(treti_trida$body,treti_trida$matem,sd)
      dobra prumerna      slaba
9.381535 10.594323  5.824824

> library(DescTools)
> LeveneTest(treti_trida$body,treti_trida$matem)
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group 2  1.1621 0.3172
      95

> vystup<-aov(treti_trida$body~treti_trida$matem)
> anova(vystup)
Analysis of Variance Table

Response: tret_i_trida$body
      Df Sum Sq Mean Sq F value    Pr(>F)
tret_i_trida$matem  2 2236.6 1118.30  11.965 2.32e-05 ***
Residuals          95 8879.0   93.46
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
> par(mfrow=c(2,2))
> plot(vystup)
```



```
> PostHocTest(vystup,method=c('hsd'))
```

```
Posthoc multiple comparisons of means : Tukey HSD  
95% family-wise confidence level
```

```
$`treti_trida$matem`
```

	diff	lwr.ci	upr.ci	pval	
prumerna-dobra	-6.551619	-11.46419	-1.6390528	0.0057	**
slaba-dobra	-16.038462	-24.78044	-7.2964852	9.4e-05	***
slaba-prumerna	-9.486842	-18.44096	-0.5327246	0.0352	*

```
---
```

```
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```