**ΑΝΑΛΥΣΗ ΔΙΑΚΥΜΑΝΣΗΣ - ΑΣΚΉΣΕΙΣ**

**Άσκηση 1.** Τα παρακάτω δεδομένα προέρχονται από μετρήσεις του δείκτη του σακχάρου στο αίμα 30 ποντικών που εξετάσθηκαν: 1) υπό κανονικές συνθήκες, 2) μετά από ένεση pitressin, 3) μετά από ένεση pitocin (στα ποντίκια δεν επιδρούν άλλοι παράγοντες π.χ. ηλικία, βάρος, φύλο κ.τ.λ.)

| | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 93 | 111 | 118 | 98 | 114 | 107 | 108 | 120 | 106 | 96 | 114 | 125 |
| **2** | 134 | 129 | 121 | 127 | 138 | 126 | 134 | 130 | 147 | 135 | | |
| **3** | 118 | 106 | 115 | 95 | 109 | 108 | 104 | 108 | | | | |

Για να εισάγουμε τα δεδομένα για ανάλυση:

```
x1<-c(93, 111, 118, 98, 114, 107, 108, 120, 106, 96, 114, 125)
y1<-c(134, 129, 121, 127, 138, 126, 134, 130, 147, 135)
z1<-c(118, 106, 115, 95, 109, 108, 104,108)
```

```
factor.level<-c(rep(1, times=length(x1)), rep(2, times=length(y1)), rep(3, 
times=length(z1)))
```

```
df.1<-data.frame(data=c(x1, y1, z1), group=factor(factor.level, 
labels=c("1","2","3")))
```

1) Κατασκευάστε ένα Boxplot των δεδομένων.

```
boxplot(data~group, data=df.1) # boxplot
```

2) Εκτιμήστε τα $\mu_1$, $\mu_2$, $\mu_3$ σημειακά και με δ.ε. 95%

```
fit1<-lm(data~group, data=df.1)
```

```
summary(fit1) # εκτίμηση παραμέτρων του μοντέλου Y_{ij} = μ_i + ε_{ij}, i=1,2,3, 
j=1,..,30.
```

```
confint.lm(fit1) #διαστήματα εμπιστοσύνης
```

3) Εξετάστε την υπόθεση $H_0$: $\mu_1 = \mu_2 = \mu_3$ με $H_1$: όχι $\mu_1 = \mu_2 = \mu_3$ σε ε.σ. $\alpha = 1\%$.

```
summary(aov(data ~ group, data = df.1)) #F test για τον έλεγχο της H_0
pairwise.t.test(df.1$data, df.1$group, data=df.1) # ανά 2 έλεγχοι ισότητας μέσων
```

4) Να δοθούν ταυτόχρονα δ.ε. για τις διαφορές $\mu_2 - \mu_1, \mu_3 - \mu_1, \mu_3 - \mu_2$

```
confuse<-TukeyHSD(m1)
```

**Άσκηση 2.** Εκλέχθηκαν τυχαία 50 δείγματα από πέντε τύπους αυτοκινήτων (1,2,3,4,5) και μετρήθηκε η κατανάλωση βενζίνης σε κάθε περίπτωση (κάτω από τις ίδιες συνθήκες κυκλοφορίας). Τα αποτελέσματα (σε χιλιόμετρα ανά γαλόνι) ήταν:

| **τύπος** | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **1** | 13,0 | 15,2 | 21,2 | 21,4 | 21,7 | 17,7 | 19,3 | 16,5 | 16,1 | 14,6 | 17,7 | | |
| **2** | 25,4 | 27,4 | 29,1 | 30,1 | 25,6 | 23,5 | 29,2 | 29,8 | | | | | |
| **3** | 16,6 | 23,0 | 20,6 | 19,9 | 22,0 | 25,2 | 16,7 | 25,2 | 24,5 | 27,1 | | | |
| **4** | 27,6 | 24,4 | 28,9 | 32,8 | 20,3 | 27,8 | 22,4 | 23,6 | | | | | |
| **5** | 31,8 | 32,3 | 31,1 | 27,9 | 30,2 | 35,7 | 32,1 | 38,5 | 30,6 | 28,5 | 29,6 | 34,9 | 36,6 |

1) Κατασκευάστε ένα Boxplot των δεδομένων.

2) Εκτιμήστε τα $\mu_1,\ldots,\mu_5$ σημειακά και με δ.ε. 95%.
3) Εξετάστε την υπόθεση $H_0$: $\mu_1 = \ldots = \mu_5$ με $H_1$: όχι $\mu_1 = \ldots = \mu_5$ σε ε.σ. α = 1%.
4) Βρείτε 95% δ.ε. για τη διαφορά $\mu_5 - \mu_1$ και 95% δ.ε. για τη διαφορά $\mu_4 - \mu_2$ και εξετάστε αν $\mu_4 = \mu_2$ (με εναλλακτική $\mu_4 \neq \mu_2$) σε ε.σ. 1%.
5) Να δοθούν ταυτόχρονα δ.ε. για όλες τις ανά δύο διαφορές $\mu_i - \mu_j$ με τη μέθοδο Tukey. Κατατάξτε τα $\mu_1, \mu_2, \ldots, \mu_5$ με πιθανότητα σφάλματος 5%.