- 2.5 – **[8 pts]**
  1. The individuals appear to be independent among groups (i.e., an individual cannot possibly be in more than one group as that person cannot have two reasons for the syndrome). They are likely independent within groups but it is not obviously so with the information given.
  2. The Levene's test suggests that the variances are equal (p=0.2181). The residual plot **(Figure B.1)** is not of much help because individual 20 appears to be a large outlier.

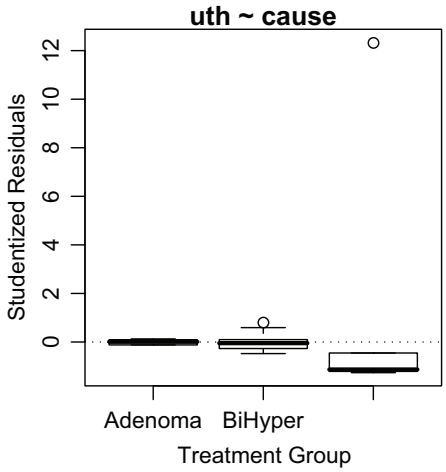

Figure B.1: Residual plot from the one-way ANOVA of UTH levels by syndrome type.

  3. The Anderson-Darling normality test strongly suggests that the residuals are not normally distributed (p<0.00005). The histogram of the residuals **(Figure B.2)** is not of much help because of a very large individual.

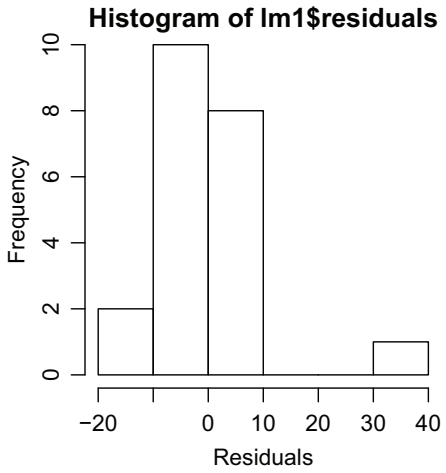

Figure B.2: Histograms of residuals from the one-way ANOVA of UTH levels by syndrome type.

  4. Observation 20 appears to be a significant outlier (p<0.00005). This was also evident on the residual plot **(Figure B.1)**.

R commands

```
> cause <- c(rep("Adenoma", 6), rep("BiHyper", 10), rep("Carcinoma", 5))
> cause <- factor(cause, levels = c("Adenoma", "BiHyper", "Carcinoma"))
> uth <- c(3.1, 3, 1.9, 3.8, 4.1, 1.9, 8.3, 3.8, 3.9, 7.8, 9.1, 15.4,
+     7.7, 6.5, 5.7, 13.6, 10.2, 9.2, 9.6, 53.8, 15.8)
> d <- data.frame(uth, cause)
> attach(d)
> lm1 <- lm(uth ~ cause)
> levene.test(lm1)
> residual.plot(lm1)
> ad.test(lm1$residuals)
> hist(lm1$residuals, xlab = "Residuals")
> outlier.test(lm1)
> detach(d)
```