> summary(model\_lm)

Call:

lm(formula = FG ~ ., data = bloodtest)

Residuals:

Min 1Q Median 3Q Max

-2.1493 -0.8115 -0.1042 0.6003 3.3438

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 2.739111 0.852417 3.213 0.00148 \*\*

TG -0.154495 0.069815 -2.213 0.02779 \*

TC 0.008878 0.037991 0.234 0.81542

`HDL-C` 0.191392 0.090579 2.113 0.03558 \*

`LDL-C` 0.169340 0.102243 1.656 0.09891 .

SBP 0.025106 0.004731 5.306 2.44e-07 \*\*\*

DBP -0.010273 0.009816 -1.047 0.29632

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 1.094 on 254 degrees of freedom

(1 observation deleted due to missingness)

Multiple R-squared: 0.1451, Adjusted R-squared: 0.1249

F-statistic: 7.187 on 6 and 254 DF, p-value: 4.277e-07

首先看模型的p值，小于0.05，说明自变量和因变量是有显著统计意义的。

再看自变量的p值，数值越小说明对因变量（血糖FG）的影响越有显著统计意义，我们可以看出收缩压SBP，高密度脂蛋白胆固醇HDL-C对血糖TG影响大，而其他自变量影响较弱。

最后看模型的R平方， R平方处于0-1之间，越靠近1说明更多的数据可以被这个模型解释，越靠近0说明越少的数据被这个模型解释。

> tidy(model\_lm)

# A tibble: 7 × 5

term estimate std.error statistic p.value

<chr> <dbl> <dbl> <dbl> <dbl>

1 (Intercept) 2.74 0.852 3.21 0.00148

2 TG -0.154 0.0698 -2.21 0.0278

3 TC 0.00888 0.0380 0.234 0.815

4 `HDL-C` 0.191 0.0906 2.11 0.0356

5 `LDL-C` 0.169 0.102 1.66 0.0989

6 SBP 0.0251 0.00473 5.31 0.000000244

7 DBP -0.0103 0.00982 -1.05 0.296

单独取出模型参数，可以写成公式：

患者的：

FG 血糖 = 2.74 - 0.154 TG甘油三酯 + 0.00888TC总胆固醇 + 0.191HDL-C高密度脂蛋白胆固醇

+ 0.169LDL-C低密度脂蛋白胆固醇 + 0.0251SBP收缩压 – 0.0103DBP舒张压

> confint(model\_lm)

2.5 % 97.5 %

(Intercept) 1.06040540 4.417816270

TG -0.29198529 -0.017004078

TC -0.06593983 0.083695290

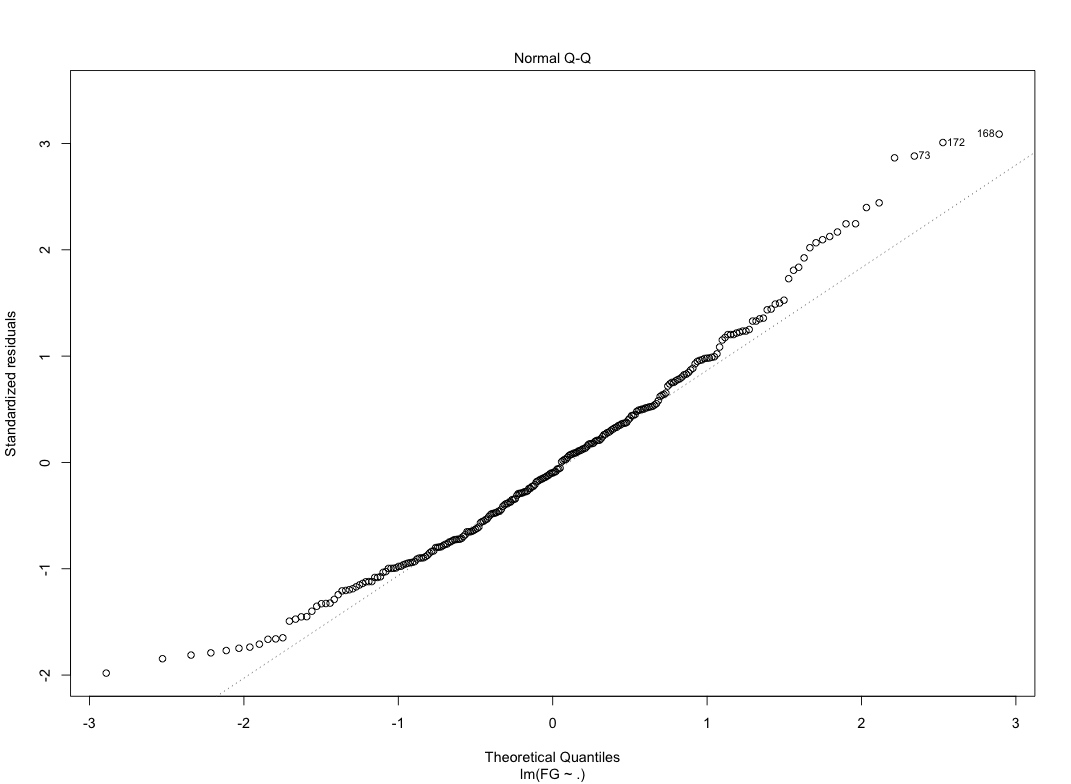
`HDL-C` 0.01300976 0.369773252

`LDL-C` -0.03201227 0.370692301

SBP 0.01578834 0.034423945

DBP -0.02960409 0.009058777

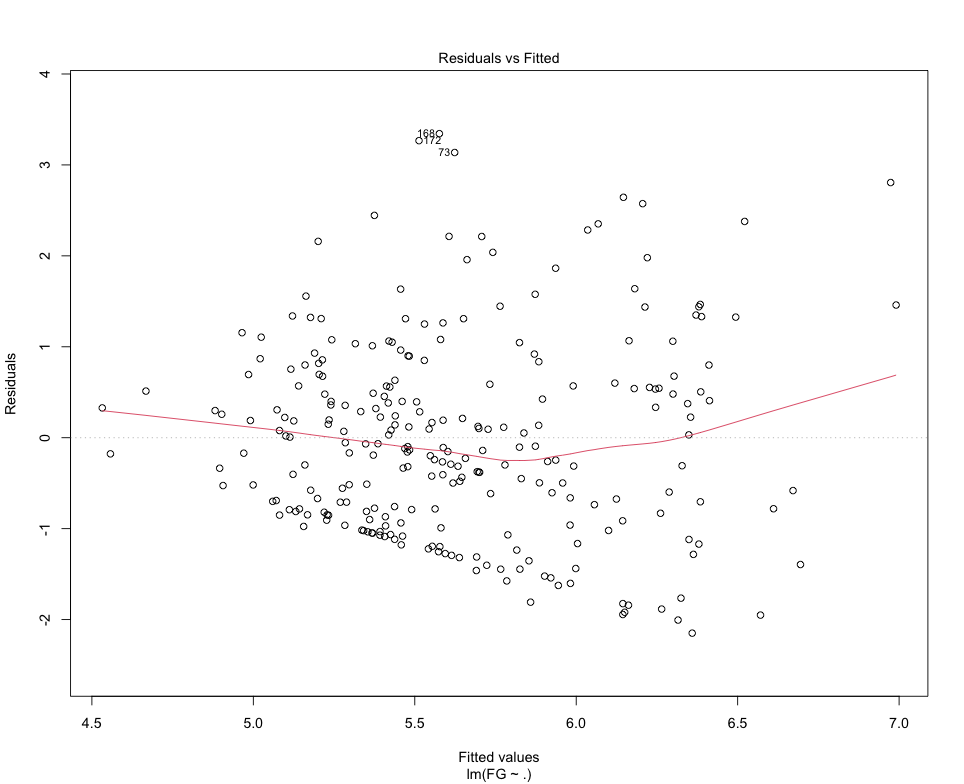
单独提取出每个自变量的95%置信区间



Normality of Residuals：

检验残差值是否正态分布。如果残差值正态分布，图像大致分散在一条直线上。

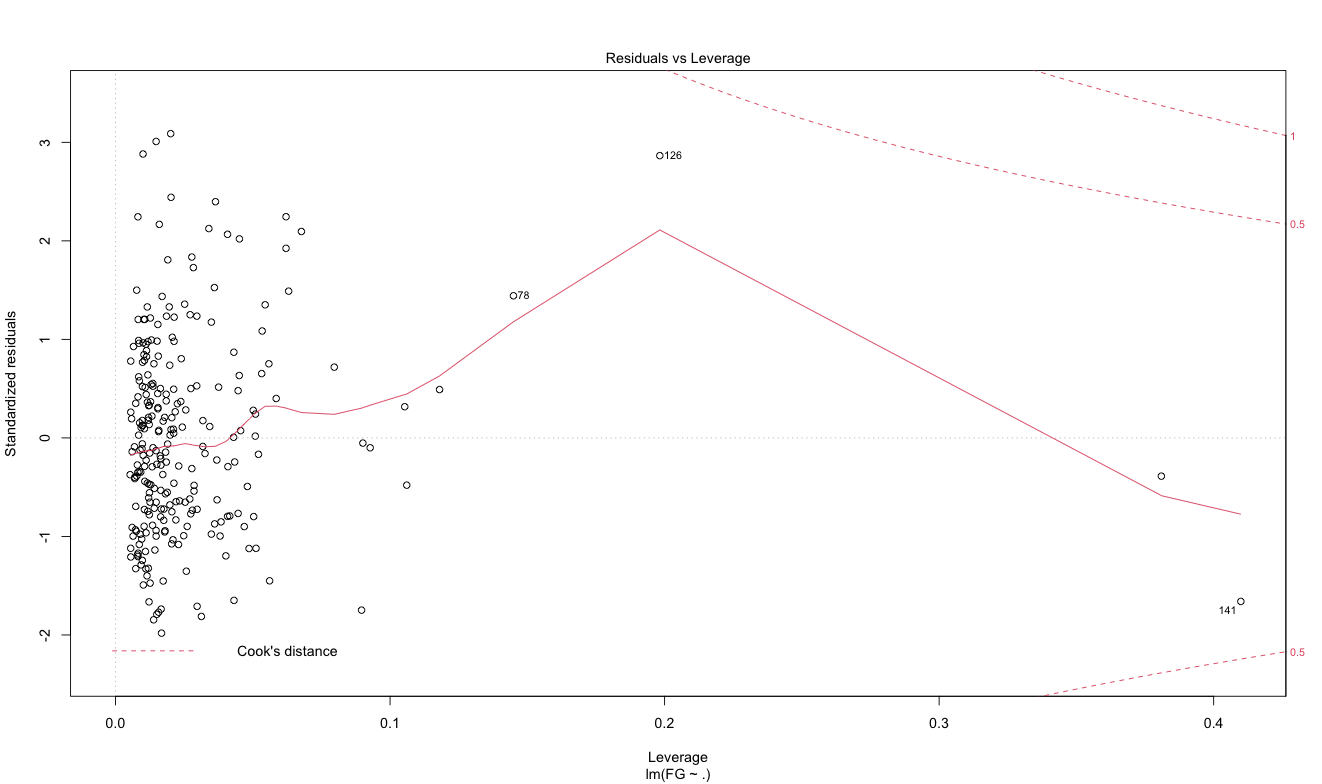
从上图看出，绝大部分残差值正态分布，右上角第73，172，168号数据的残差值稍有偏差。



Linearity of the data:

检验原始数据是否有线性关系。最佳状态红色线条呈水平状态。

从上图看出，红色线条接但不是完全水平状态，说明模型的线性关系稍有偏差。



Outliers and high leverage points：

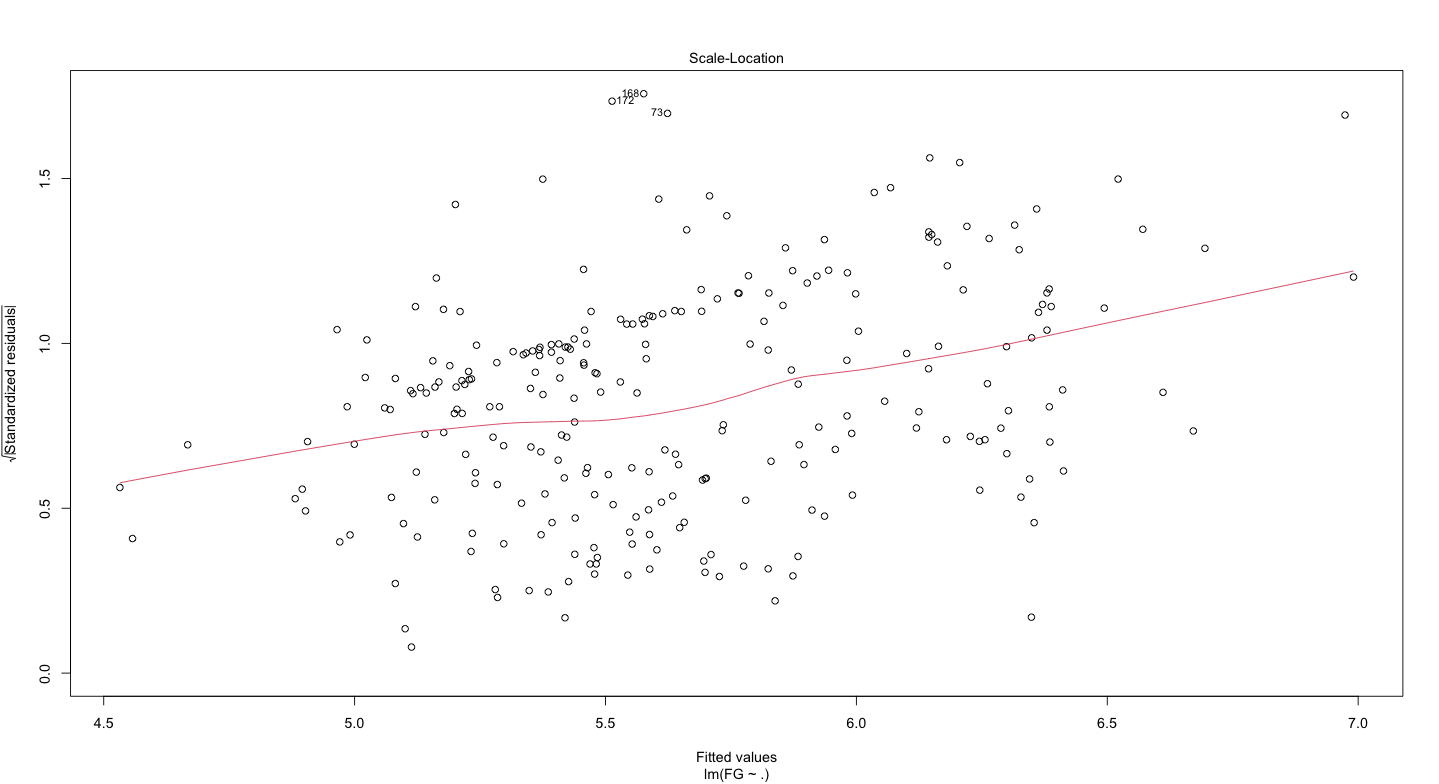
检验原始数据是否有极值和high leverage point（不理想的自变量）

从上图看出，原始数据有少数极值（纵轴绝对值大于3），

standardized residuals（标准化的残差值）（红色线条）受78，126，141号数据影响较大，理想状态应该为水平状态。

检验统计值 2 \*（p+1) / n = 2 \* (自变量数目+1) / 观察数目 = 2 \* （6+1）/ 262 = 14/262 = 0.053

从上图看出，原始数据都呈现在统计检验值之下（横轴小于0.053），是理想的



Homogeneity of Variance：

检验残差值是否有固定的方差。理想状态图像沿着水平的红线均匀分布。

从上图看出，这个模型的残差值没有理想的固定方差，而是呈异方差性（heteroscedasticity）。