Using R and the `anova` function we can easily compare **nested** models. Where we are dealing with regression models, then we apply the `F-Test` and where we are dealing with logistic regression models, then we apply the `Chi-Square Test`. By nested, we mean that the independent variables of the **simple** model will be a **subset** of the more **complex** model. In essence, we try to find the best parsimonious fit of the data. Note that we should fit the models on the same dataset.

The Null Hypothesis is that the **simple** model is better and we reject the null hypothesis if the p-value is less than 5% inferring that the **complex** model is is significantly better than the simple one.

Example of Comparing Nested Models

Let's work with the `LifeCycleSavings` dataset by considering as dependent variable the `sr` and the rest as independent variables (IV).

```
> LifeCycleSavings
                  sr pop15 pop75     dpi  ddpi
Australia      11.43 29.35  2.87 2329.68  2.87
Austria        12.07 23.32  4.41 1507.99  3.93
Belgium        13.17 23.80  4.43 2108.47  3.82
Bolivia         5.75 41.89  1.67  189.13  0.22
Brazil         12.88 42.19  0.83  728.47  4.56
Canada          8.79 31.72  2.85 2982.88  2.43
Chile           0.60 39.74  1.34  662.86  2.67
China          11.90 44.75  0.67  289.52  6.51
Colombia        4.98 46.64  1.06  276.65  3.08
Costa Rica     10.78 47.64  1.14  471.24  2.80
Denmark        16.85 24.42  3.93 2496.53  3.99
Ecuador         3.59 46.31  1.19  287.77  2.19
Finland        11.24 27.84  2.37 1681.25  4.32
France         12.64 25.06  4.70 2213.82  4.52
Germany        12.55 23.31  3.35 2457.12  3.44
Greece         10.67 25.62  3.10  870.85  6.28
Guatamala       3.01 46.05  0.87  289.71  1.48
Honduras        7.70 47.32  0.58  232.44  3.19
```

Let's say that we can to compare the following two models:

- `fit0` which is the \(sr = \alpha\) **VS**
- `fit1` which is the \(sr = \alpha +\beta \times pop15\)

```
fit0 <- lm(sr ~ 1, data = LifeCycleSavings)

fit1 <- lm(sr ~ pop15, data = LifeCycleSavings)


summary(fit0)
summary(fit1)
```

```
> summary(fit0)

Call:
lm(formula = sr ~ 1, data = LifeCycleSavings)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.071 -2.701  0.839  2.946 11.429 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)   9.6710     0.6336   15.26   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.48 on 49 degrees of freedom
```

```
> summary(fit1)

Call:
lm(formula = sr ~ pop15, data = LifeCycleSavings)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.637 -2.374  0.349  2.022 11.155 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept) 17.49660    2.27972   7.675 6.85e-10 ***
pop15       -0.22302    0.06291  -3.545 0.000887 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 4.03 on 48 degrees of freedom
Multiple R-squared:  0.2075,	Adjusted R-squared:  0.191 
F-statistic: 12.57 on 1 and 48 DF,  p-value: 0.0008866
```

Notice the the P-value of the F-Test of the `fit1` model is **0.0008866** which actually actually tests the Null Hypothesis that “all the beta coefficients are zero” versus the alternative hypothesis that “at least one beta coefficient is not zero”. Since we have only one beta coefficient, the `pop15` the p-value of the F-Test is the same with the p-value of the T-Test as we can see above.

Now, if we compare the `fit0` vs the `fit1`, in essence, we test if we should include the `pop15` coefficient or not, thus we expect to get the same p-value. Let’s compare the nested models using `anova`:

```
anova(fit0, fit1, test='F')
```

```
> anova(fit0, fit1, test='F')
Analysis of Variance Table

Model 1: sr ~ 1
Model 2: sr ~ pop15
  Res.Df    RSS Df Sum of Sq      F    Pr(>F)    
1     49 983.63                                  
2     48 779.51  1    204.12 12.569 0.0008866 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

As, expected we got the same p-value, and we can say that we should prefer the `fit1` compared to `fit0` model.

Let’s make another comparison by comparing the `fit1` compared to the `fit4` which contains all the IVs.

```
fit4<-lm(sr~pop15+pop75+dpi+ddpi, data = LifeCycleSavings)

summary(fit4)
```

```
> summary(fit4)

Call:
lm(formula = sr ~ pop15 + pop75 + dpi + ddpi, data = LifeCycleSavings)

Residuals:
    Min      1Q  Median      3Q     Max 
-8.2422 -2.6857 -0.2488  2.4280  9.7509 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 28.5660865  7.3545161   3.884 0.000334 ***
pop15       -0.4611931  0.1446422  -3.189 0.002603 ** 
pop75       -1.6914977  1.0835989  -1.561 0.125530    
dpi         -0.0003369  0.0009311  -0.362 0.719173    
ddpi         0.4096949  0.1961971   2.088 0.042471 *  
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.803 on 45 degrees of freedom
Multiple R-squared:  0.3385,	Adjusted R-squared:  0.2797 
F-statistic: 5.756 on 4 and 45 DF,  p-value: 0.0007904
```

Let’s compare the two models:

```
anova(fit1, fit4, test='F')
```

```
> anova(fit1, fit4, test='F')
Analysis of Variance Table

Model 1: sr ~ pop15
Model 2: sr ~ pop15 + pop75 + dpi + ddpi
  Res.Df    RSS Df Sum of Sq     F  Pr(>F)  
1     48 779.51                             
2     45 650.71  3     128.8 2.969 0.04177 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The p-value is **0.04177** forcing us to reject the null hypothesis that the `fit1` models is better. Finally, let’s compare the `fit1` model versus the `fit3` which contains the first 3 IV of the dataset.

```
fit3<-lm(sr~pop15+pop75+dpi, data = LifeCycleSavings)
anova(fit1, fit3, test='F')
```

```
> anova(fit1, fit3, test='F')
Analysis of Variance Table

Model 1: sr ~ pop15
Model 2: sr ~ pop15 + pop75 + dpi
  Res.Df    RSS Df Sum of Sq      F Pr(>F)
1     48 779.51                           
2     46 713.77  2    65.744 2.1185 0.1318
```

In this case, the p-value is **0.1318** which means that we should accept the null hypothesis that the `fit1` is better than the `fit3`.