

## #INTERPRETATION DES TEST DE FISHER ET STUDENT SUR LE MODELE 1

```
Residuals:
    Min       1Q   Median       3Q      Max
-5.3396 -2.1501 -0.9674  1.1921 16.6085

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -0.90485    0.68497  -1.321   0.187
educ         0.54136    0.05325  10.167 <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.378 on 524 degrees of freedom
Multiple R-squared:  0.1648,    Adjusted R-squared:  0.1632
F-statistic: 103.4 on 1 and 524 DF,  p-value: < 2.2e-16
```

Dans la sortie que vous avez fournie, les résultats du test de Fisher (Test F) et du test de Student pour la variable 'educ' sont inclus. Voici comment les interpréter :

### 1. Test de Fisher (Statistique F) :

- La statistique F est présentée sous la section "F-statistic" dans la sortie du modèle. Dans cet exemple, la statistique F est de 103.4.
- Les degrés de liberté associés sont indiqués comme "1 and 524 DF" (DF signifie "degrés de liberté").
- La p-valeur associée à la statistique F est très faible,  $< 2.2e-16$ , indiquant que le modèle dans son ensemble est statistiquement significatif.

*Interprétation :* Le modèle dans son ensemble fournit une explication significative de la variance de la variable dépendante.

### 2. Tests de Student (t-tests) pour 'educ' :

- Les résultats pour 'educ' sont présentés sous la section "Coefficients".
- La statistique t associée à 'educ' est de 10.167.
- La p-valeur associée à 'educ' est indiquée comme " $< 2e-16$ ", ce qui est une manière abrégée d'écrire une p-valeur très faible.

*Interprétation :* Le coefficient associé à 'educ' est significativement différent de zéro. En d'autres termes, le niveau d'éducation (variable 'educ') a un effet significatif sur la variable dépendante.

En résumé, les résultats suggèrent que le modèle dans son ensemble est significatif, et la variable 'educ' a un effet significatif sur la variable dépendante, selon le test de Fisher et les tests de Student associés. Les "Signif. codes" dans la sortie indiquent également le niveau de signification des p-valeurs, où "\*\*\*\*" signifie une p-valeur très faible ( $< 0.001$ ).

# TEST DE CHOW

R 4.3.1 · C:/Users/bauer/OneDrive/Bureau/économétrie/Projet CDK/ ↗

## Residuals:

Min	1Q	Median	3Q	Max
-5.5532	-1.9801	-0.7071	1.2030	15.8370

## Coefficients:

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	-3.39054	0.76657	-4.423	1.18e-05 ***
educ	0.64427	0.05381	11.974	< 2e-16 ***
exper	0.07010	0.01098	6.385	3.78e-10 ***

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

Residual standard error: 3.257 on 523 degrees of freedom  
Multiple R-squared: 0.2252, Adjusted R-squared: 0.2222  
F-statistic: 75.99 on 2 and 523 DF, p-value: < 2.2e-16

```
# Supposons que votre modèle complet s'appelle 'modele_complet'
# Supposons que vous avez deux sous-groupes 'sous_groupe_educ' et 'sous_groupe_exper'

# Estimation des modèles pour chaque sous-groupe
modele_educ <- lm(y ~ educ, data = sous_groupe_educ)
modele_exper <- lm(y ~ exper, data = sous_groupe_exper)

# Estimation du modèle global
modele_complet <- lm(y ~ educ + exper, data = toutes_les_observations)

# Calcul de la statistique de test de Chow
SSR_complet <- sum(residuals(modele_complet)^2)
SSR_educ <- sum(residuals(modele_educ)^2)
SSR_exper <- sum(residuals(modele_exper)^2)

# Nombre de sous-groupes
k <- 2

# Degrés de liberté des résidus dans le modèle complet
df_complet <- nrow(toutes_les_observations) - length(coef(modele_complet))

# Statistique de test de Chow
stat_chow <- ((SSR_complet - (SSR_educ + SSR_exper)) / k) / ((SSR_educ + SSR_exper) / df_complet)
```

La statistique de test de Chow suit une distribution de probabilité F. Vous pouvez comparer la statistique de test calculée avec les valeurs critiques de la distribution F pour déterminer la significativité du test. Une p-valeur associée à la statistique de test de Chow inférieure à un seuil de signification (comme 0.05) suggère le rejet de l'hypothèse nulle d'égalité des coefficients entre les sous-groupes.

En fonction des résultats, vous pourrez interpréter si les coefficients de 'educ' et 'exper' diffèrent significativement entre les deux sous-groupes.

## INTERPRETATION DU COEFFICIENT DE DETERMINATION $R^2$

Le coefficient de détermination, souvent appelé  $R^2$ , est une mesure importante dans la régression linéaire en économétrie.  $R^2$  quantifie la proportion de la variabilité totale de la variable dépendante ( $Y$ ) qui est expliquée par le modèle. Voici comment interpréter  $R^2$  :

- $R^2$  varie de 0 à 1. Plus il est proche de 1, plus le modèle explique une grande partie de la variabilité de la variable dépendante.
- $R^2$  égal à 0 signifie que le modèle n'explique aucune variabilité et que la variable dépendante est totalement imprévisible à partir des variables indépendantes du modèle.
- $R^2$  égal à 1 signifierait que le modèle explique parfaitement toute la variabilité de la variable dépendante.
- $R^2$  peut également être interprété comme la proportion de la somme des carrés expliqués ( $SCE$ ) par rapport à la somme totale des carrés ( $SCT$ ). Plus précisément,  $R^2 = \frac{SCE}{SCT}$ .

Cependant, il est important de noter que  $R^2$  seul ne donne pas d'informations sur la validité des coefficients individuels du modèle. Un  $R^2$  élevé ne garantit pas que chaque coefficient est statistiquement significatif ni que le modèle est bien spécifié.

En résumé, un  $R^2$  élevé suggère que le modèle explique une grande partie de la variabilité de la variable dépendante, mais pour une évaluation complète, d'autres diagnostics du modèle doivent également être pris en compte.

# INTERPRETATION GENERALE MODELE ECONOMETRIQUE

## Comparaison des Modèles

Dependent variable:			
	(1)	wage (2)	(3)
educ	0.541*** (0.053)	0.644*** (0.054)	0.599*** (0.051)
exper		0.070*** (0.011)	0.022* (0.012)
Constant	-0.905 (0.685)	-3.391*** (0.767)	-2.873*** (0.729)
Observations	526	526	526
R2	0.165	0.225	0.306
Adjusted R2	0.163	0.222	0.302
Residual Std. Error	3.378 (df = 524)	3.257 (df = 523)	3.084 (df = 522)
F Statistic	103.363*** (df = 1; 524)	75.990*** (df = 2; 523)	76.873*** (df = 3; 522)

Note: \*p<0.1; \*\*p<0.05; \*\*\*p<0.01

Error in if (nchar(text.matrix[r, c]) > max.length[real.c]) { :  
missing value where TRUE/FALSE needed

### 1. Coefficients de Régression :

- Model 1 (educ):** Le coefficient est 0.541 avec une p-valeur très faible (<0.01). Cela suggère que le niveau d'éducation est significativement lié au salaire.
- Model 2 (educ, exper):** L'ajout de l'expérience (exper) a augmenté le coefficient de l'éducation (educ) à 0.644. L'expérience a également un effet significatif sur le salaire.

### 2. Constante :

- Model 1 :** La constante est -0.905.
- Model 2 :** La constante a changé à -3.391 après l'ajout de l'expérience.

### 3. R2 et Adjusted R2 :

- Model 1 :** R2 = 0.165, Adjusted R2 = 0.163.
- Model 2 :** R2 = 0.225, Adjusted R2 = 0.222.
- Model 3 :** R2 = 0.306, Adjusted R2 = 0.302.

Les valeurs R2 indiquent la proportion de variance expliquée, et Adjusted R2 ajuste pour le nombre de variables. Les modèles successifs semblent expliquer davantage de variance.

### 4. Test F :

- Model 1 vs. Model 2 :** F Statistic = 103.363 avec une p-valeur très faible (<0.01). Le modèle global avec educ et exper est statistiquement significatif par rapport au modèle avec seulement educ.
- Model 2 vs. Model 3 :** F Statistic = 75.990 avec une p-valeur très faible (<0.01). L'ajout de la tenure est significatif.
- Model 3 vs. Model 4 :** F Statistic = 76.873 avec une p-valeur très faible (<0.01). L'ajout de female est significatif.

En résumé, l'ajout d'expérience, tenure et female aux modèles successifs semble améliorer significativement l'ajustement des modèles, ce qui est reflété par les augmentations de R2 et les tests F significatifs. Les coefficients de régression et les p-valeurs fournissent des informations sur la force et la signification de chaque variable dans le modèle.

