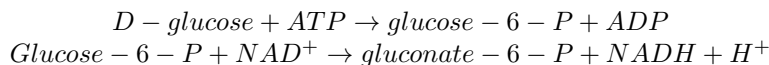


Séance 2 : Régression linéaire avec R - courbe d'étalonnage

Exercice 0.1 Etalonnage pour le dosage du glucose

Pour pouvoir déterminer la concentration du Glucose à partir de la mesure de l'absorbance, on effectue une calibration préliminaire: une courbe d'étalonnage du dosage du Glucose. Le principe du dosage est basé sur les réactions suivantes



On suppose que la formation du Glucose est donc proportionnelle à celle du NADH dont on va mesurer l'apparition avec le spectrophotomètre. Il nous faut donc connaître le coefficient de proportionnalité qui permet à partir de l'absorbance de déterminer la concentration en Glucose.

Une courbe étalon est donc réalisée à partir d'une solution mère que l'on dilue et dont on mesure l'absorbance: on obtient ainsi un nuage de points $(C_{\text{Glc},i}, \text{absorbance}_i)$

Solution	[glucose] (g/L)	Absorbance
Blanc	0	0
Dilution 1/4	0.123	0.1
Dilution 1/2	0.246	0.199
Dilution 3/4	0.369	0.301
Non diluée	0.492	0.411

1. Lire le fichier Etalon.csv à l'aide de l'instruction read.csv2. On appellera deta la data frame et on fera un attach(deta) pour simplifier le nom des variables.
2. Afficher le nuage de points $(\text{glucose}_i, \text{absorbance}_i)$.
3. Déterminer l'équation de la droite de régression entre concentration de Glucose et absorbance à l'aide de la fonction "lm". On appellera reg le résultat de lm. A quoi correspond l'instruction coef(reg) ? Donner l'équation de la droite de régression
4. Afficher le nuage de points, la droite de régression sur un même graphe. Mettre à ce graphique un titre principal "Etalonnage Glucose vs absorbance" puis l'équation de la droite de régression et le coefficient de détermination R^2 en sous-titre.

Le R^2 (la variance expliquée par le modèle divisée par la variance totale) étant très proche de 1, cela signifie que le pouvoir prédictif du modèle est fort.

5. Etude de la qualité de la régression.

(a) A partir de la commande summary(reg), remplir le tableau suivant

Coefficients	Estimate	Std. Error	t value	Pr(> t)
(Intercept)				
x				

(b) Quels sont les estimations des écarts types des coefficients de la droite ?

Pour chaque paramètre, la table donne la statistique observée (colonne t value) ainsi que la pvalue (colonne $Pr(>|t|)$) associée au test d'hypothèse H_0 : "le paramètre est nul" contre H_1 : "le paramètre n'est pas nul". Peut-on accepter l'hypothèse "la courbe passe par l'origine" ? Si oui, tester un autre modèle plus adapté.

6. Etudier les résidus

7. Prédiction

(a) Prédire les valeurs suivantes

<i>[glucose] (g/L)</i>	<i>Absorbance</i>	<i>Absorbance min</i>	<i>Absorbance max</i>
0.05			
0.1			
0.15			
0.2			
0.25			
0.3			
0.35			
0.4			

(b) Tracer la droite de régression ainsi que les intervalles de confiance et de prédiction (bornes min et max).

8. Calibration: courbe d'étalonnage

On n'a pas fait une régression linéaire glucose (en ordonnée) en fonction de l'absorbance (en abscisse) puisque c'est la concentration de glucose qui est connue (variable explicative supposée certaine), l'absorbance mesurée étant l'estimation d'une variable aléatoire. Mais ensuite, c'est la fonction réciproque qui nous intéresse: connaissant la mesure de l'absorbance, on veut en déduire la concentration (inconnue) en glucose d'un soluté qu'on passera au spectromètre. On ne va pas donc pas utiliser la fonction `predict` mais la fonction `calibrate` du package `investr`. Cette fonction nous donnera à partir d'une absorbance mesurée la concentration de glucose correspondante et un intervalle de prédiction.

(a) Charger le package `investr`. On a obtenu une absorbance de 0.2. Quelle est la concentration de glucose correspondante ?

```
> library(investr)
> # absorbance de 0.20 donne concentration glc et un IC dans res
> res <- calibrate(reg, y0 = 0.2, interval = "inversion", level = 0.9)
> print(res)
> plotFit(reg, interval = "prediction", level = 0.9, shade = TRUE, col.pred = "skyblue")
> abline(h = 0.2, v = c(res$lower, res$estimate, res$upper), lty = 2)
```

(b) On a répété trois fois la mesure et on a obtenu: 0.2; 0.22 et 0.19. Quelle est la concentration de glucose correspondante ?

```
> # si on dispose de 3 mesures de l'absorbance
> res=calibrate(reg, y0 = c(0.2,0.22,0.19), interval = "inversion", level = 0.9)
> print(res)
> plotFit(reg, interval = "prediction", level = 0.9, shade = TRUE, col.pred = "skyblue")
> abline(h = 0.2, v = c(res$lower, res$estimate, res$upper), lty = 2)
```