

Bases méthodologiques pour l'analyse de données de croissance mycélienne

Melen Leclerc

INRAE IGEPP

2025 - Rennes



Préambule

- Cours - TD initialement destiné à vous fournir les bases pour analyser les données du projet expérimental
- formalisation d'un modèle simple pour analyser la croissance d'un champignon
- analyse des données à l'aide du cadre de l'ANCOVA
- utilisation de R pour la mise en pratique

Programme

- Mercredi 05/03
- Cours TD : modèle de croissance radiale, ANCOVA
- diapos pdf, lesion_asco.csv, cours_2023.R

Programme

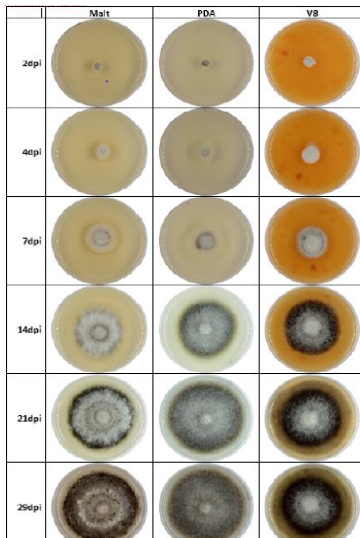
- Mercredi 05/03

- Cours TD : modèle de croissance radiale, ANCOVA
- diapos pdf, lesion_asco.csv, cours_2023.R

- Mercredi 12/03

- Soutien à l'analyse des données expérimentales, réponse aux questions
- Documents disponibles sur
https://github.com/meleninrae/cours_ancova

Suivi de croissance sur support (boîte, feuille) circulaire

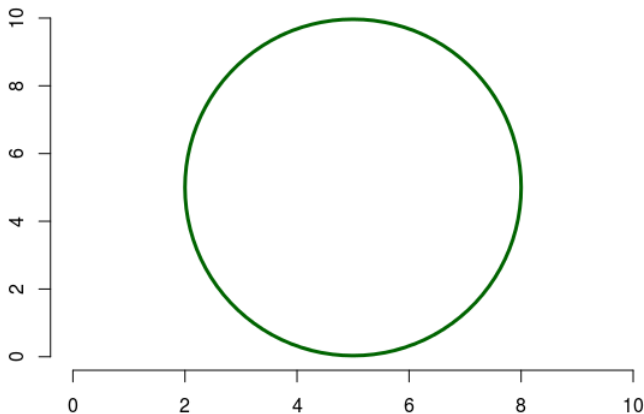


Suivi de croissance sur support (boîte, feuille) circulaire



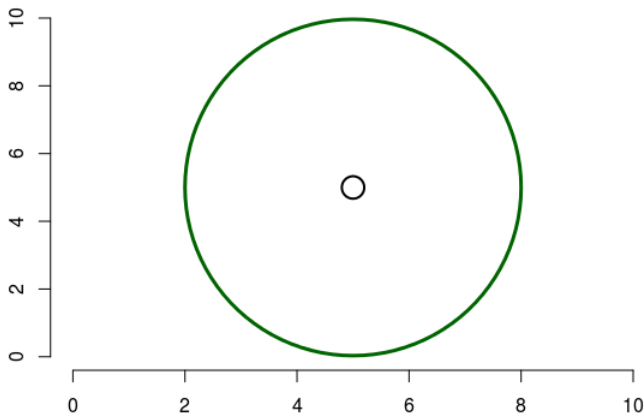
Suivi de croissance sur support (boîte, feuille) circulaire

Croissance mycélienne



Suivi de croissance sur support (boîte, feuille) circulaire

Croissance mycélienne



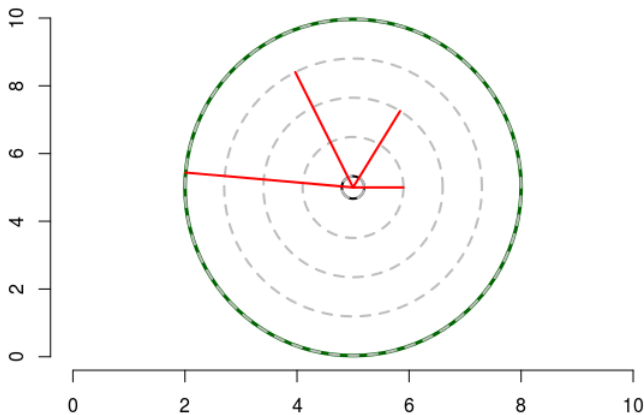
Suivi de croissance sur support (boîte, feuille) circulaire

Mesures de croissance

Suivi du rayon r au court du temps t

→ Observations : $\mathcal{Y}_i = [(r_1, t_1), \dots, (r_i, t_i), \dots, (r_N, t_N)]$

Croissance mycélienne



Modèle de croissance radiale

- ρ : vitesse radiale de développement du champignon
 - R : rayon maximal de la boîte ou de la feuille
 - r_0 : rayon initial de la colonie
- Modèle de croissance au cours du temps si vitesse constante ????

Modèle de croissance radiale

- ρ : vitesse radiale de développement du champignon
- R : rayon maximal de la boîte ou de la feuille
- r_0 : rayon initial de la colonie
- Rayon

$$r(t) = r_0 + \min(\rho t, R)$$

→ relation linéaire

- Surface

$$S(t) = \pi r(t)^2$$

→ relation quadratique

Modèle de croissance radiale

- ρ : vitesse radiale de développement du champignon
- R : rayon maximal de la boîte ou de la feuille
- r_0 : rayon initial de la colonie
- Si taille de boîte/feuille infinie (ou aucune mesure après saturation au rayon maximal)
- Rayon

$$r(t) = r_0 + \rho t$$

→ relation linéaire

Ajustement du modèle aux données

- Des observations

$$\mathcal{Y}_i = \{(r_1, t_1), \dots, (r_i, t_i), \dots, (r_N, t_N)\}$$

- Un modèle linéaire

$$r(t) = r_0 + \rho t$$

Ajustement du modèle aux données

- Des observations

$$\mathcal{Y}_i = \{(r_1, t_1), \dots, (r_i, t_i), \dots, (r_N, t_N)\}$$

- Un modèle linéaire

$$r(t) = r_0 + \rho t$$

→ Comment estimer les paramètres $\hat{\rho}$ et \hat{r}_0 ?

Ajustement du modèle aux données

- Des observations

$$\mathcal{Y}_i = \{(r_1, t_1), \dots, (r_i, t_i), \dots, (r_N, t_N)\}$$

- Un modèle linéaire

$$r(t) = r_0 + \rho t$$

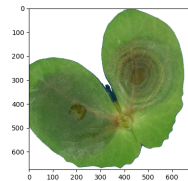
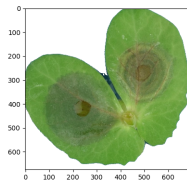
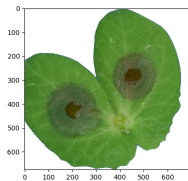
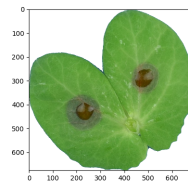
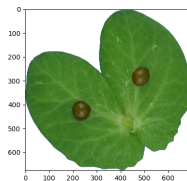
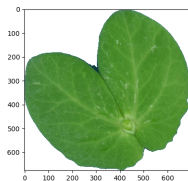
→ Comment estimer les paramètres $\hat{\rho}$ et \hat{r}_0 ?

→ Régression linéaire

$$r_i = r_0 + \rho t_i + \epsilon_i$$

Données TD

● Ascochyte du pois



Données TD

- Deux pathogènes : *Peyronellaea pinodes* (dp) & *Phoma medicaginis* (pm)
- Deux variétés-génotypes : résistant et sensible
- Données de croissance de lésion symptomatiques sur feuille

temps	genotype	espece	sfeuille	sain	lesion	rayon
Min. :1.000	resistant:5432	dp:6336	Min. : 0.888	Min. : 0.071	Min. :0.0000	Min. :0.0000
1st Qu.:2.000	sensible :5388	pm:4484	1st Qu.: 2.298	1st Qu.: 1.819	1st Qu.:0.0240	1st Qu.:0.0874
Median :4.000			Median : 3.274	Median : 2.666	Median :0.3000	Median :0.3090
Mean :4.451			Mean : 4.571	Mean : 3.600	Mean :0.9393	Mean :0.4153
3rd Qu.:7.000			3rd Qu.: 6.801	3rd Qu.: 5.823	3rd Qu.:1.6623	3rd Qu.:0.7274
Max. :9.000			Max. :10.653	Max. :10.588	Max. :5.9530	Max. :1.3766

Régression linéaire dans R

- On prend toutes les données espèce et génotype confondus

```
mod<-lm(rayon~temps,data=data)
```

Régression linéaire dans R

- On prend toutes les données espèce et génotype confondus

```
mod<-lm(rayon~temps,data=data)
```

- Résultats

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.1033394	0.0034246	-30.18	<2e-16 ***
temps	0.1165138	0.0006625	175.86	<2e-16 ***

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

Residual standard error: 0.1811 on 10818 degrees of freedom

Multiple R-squared: 0.7409, Adjusted R-squared: 0.7408

F-statistic: 3.093e+04 on 1 and 10818 DF, p-value: < 2.2e-16

Régression linéaire dans R

- On prend toutes les données espèce et génotype confondus

```
mod<-lm(rayon~temps,data=data)
```

- Résultats

Coefficients:

	Estimate	Std. Error	t value	Pr(> t)
(Intercept)	-0.1033394	0.0034246	-30.18	<2e-16 ***
temps	0.1165138	0.0006625	175.86	<2e-16 ***

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

Residual standard error: 0.1811 on 10818 degrees of freedom

Multiple R-squared: 0.7409, Adjusted R-squared: 0.7408

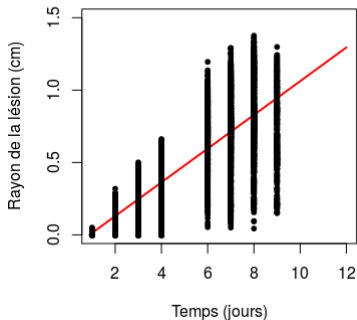
F-statistic: 3.093e+04 on 1 and 10818 DF, p-value: < 2.2e-16

- $\hat{\rho} = 0.1165$ et $\hat{r}_0 = -0.1033$

Croissance moyenne

Croissance moyenne tous effets confondus :

$$r(t) = \hat{\rho} + \hat{r}_0 t = -0.1033 + 0.1165t$$



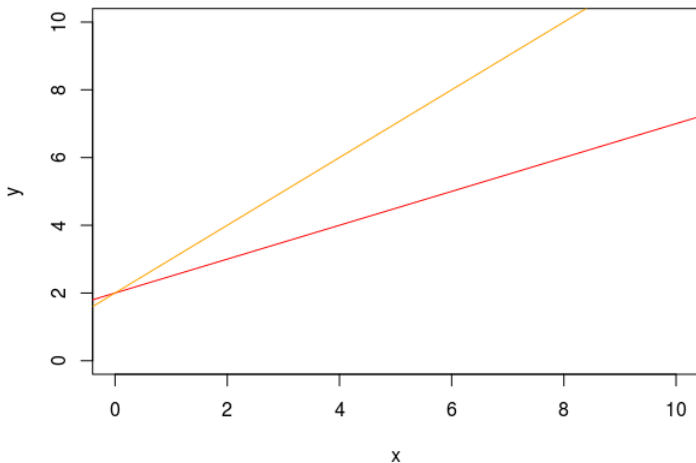
→ ordonnée à l'origine négative...

Introduction de l'effet d'une variable qualitative

- Régression linéaire : $r_i = r_0 + \rho t_i + \epsilon_i$
- Ajout d'une variable qualitative à deux modalités (e.g. effet espèce : dp & pm)
- Quels effets peut avoir cette variable qualitative sur la relation linéaire ?

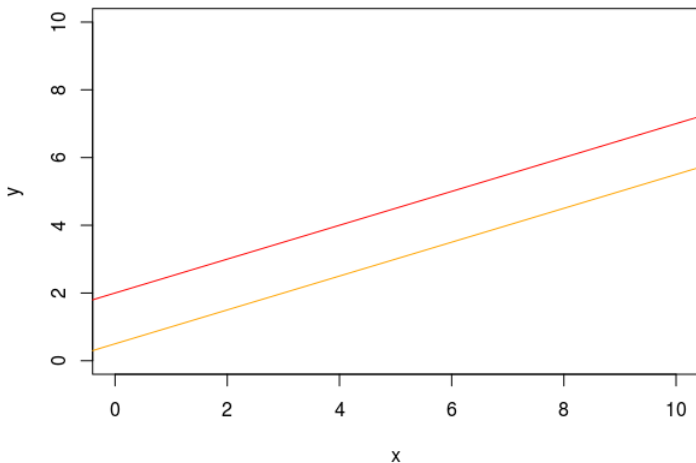
Introduction de l'effet d'une variable qualitative

- Des pentes différentes



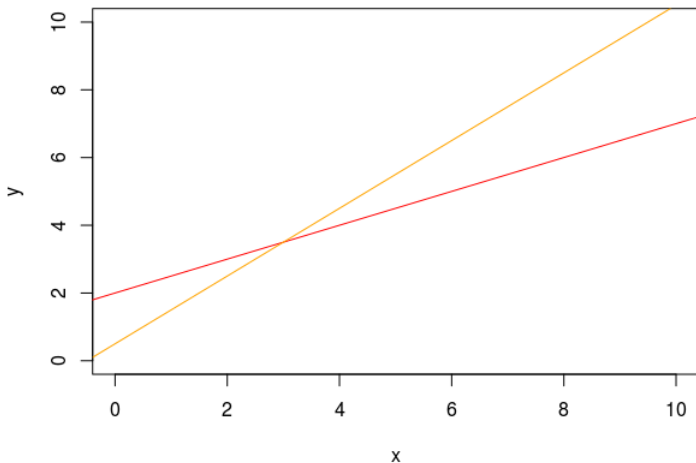
Introduction de l'effet d'une variable qualitative

- Des ordonnées à l'origine différentes



Introduction de l'effet d'une variable qualitative

- Des pentes et des ordonnées à l'origine différentes



Analyse de covariance

- Régression linéaire : $r_i = r_0 + \rho t_i + \epsilon_i$
- Variable qualitative à deux modalités

Analyse de covariance

- Régression linéaire : $r_i = r_0 + \rho t_i + \epsilon_i$
 - Variable qualitative à deux modalités
 - Idée
- ordonnées à l'origine : $\mu_{r_0} + \alpha_i$
- pente : $\mu_\rho + \beta_i$

Analyse de covariance

- Régression linéaire : $r_i = r_0 + \rho t_i + \epsilon_i$
 - Variable qualitative à deux modalités
 - Idée
- ordonnées à l'origine : $\mu_{r_0} + \alpha_i$
- pente : $\mu_\rho + \beta_i$
- Pour la modalité 1 : $r_1(t) = (\mu_{r_0} + \alpha_1) + (\mu_\rho + \beta_1)t$
 - Pour la modalité 2 : $r_2(t) = (\mu_{r_0} + \alpha_2) + (\mu_\rho + \beta_2)t$

ANCOVA et R

- Rappel modèle tous effets confondus

```
modele<-lm(rayon~temps,data=data)
```

ANCOVA et R

- Rappel modèle tous effets confondus

```
modele<-lm(rayon~temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine

ANCOVA et R

- Rappel modèle tous effets confondus

```
modele<-lm(rayon~temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine

```
modele<-lm(rayon~temps+espece,data=data)
```


ANCOVA et R

- Rappel modèle tous effets confondus

```
modele<-lm(rayon~temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine

```
modele<-lm(rayon~temps+espece,data=data)
```

- Ajout de l'effet espèce sur la pente

ANCOVA et R

- Rappel modèle tous effets confondus

```
modele<-lm(rayon~temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine

```
modele<-lm(rayon~temps+espece,data=data)
```

- Ajout de l'effet espèce sur la pente

```
modele<-lm(rayon~temps+espece:temps,data=data)
```

ANCOVA et R

- Rappel modèle tous effets confondus

```
modele<-lm(rayon~temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine

```
modele<-lm(rayon~temps+espece,data=data)
```

- Ajout de l'effet espèce sur la pente

```
modele<-lm(rayon~temps+espece:temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine et la pente

ANCOVA et R

- Rappel modèle tous effets confondus

```
modele<-lm(rayon~temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine

```
modele<-lm(rayon~temps+espece,data=data)
```

- Ajout de l'effet espèce sur la pente

```
modele<-lm(rayon~temps+espece:temps,data=data)
```

- Ajout de l'effet espèce sur l'ordonnée à l'origine et la pente

```
modele<-lm(rayon~temps+espece+espece:temps,data=data)
```

ANCOVA et R

- ANCOVA $r_i(t) = (\mu_{r_0} + \alpha_i) + (\mu_\rho + \beta_i)t + \epsilon_i$

ANCOVA et R

- ANCOVA $r_i(t) = (\mu_{r_0} + \alpha_i) + (\mu_\rho + \beta_i)t + \epsilon_i$
- $r_1(t) = (\mu_{r_0} + \alpha_1) + (\mu_\rho + \beta_1)t$
- $r_2(t) = (\mu_{r_0} + \alpha_2) + (\mu_\rho + \beta_2)t$

ANCOVA et R

- ANCOVA $r_i(t) = (\mu_{r_0} + \alpha_i) + (\mu_\rho + \beta_i)t + \epsilon_i$
- $r_1(t) = (\mu_{r_0} + \alpha_1) + (\mu_\rho + \beta_1)t$
- $r_2(t) = (\mu_{r_0} + \alpha_2) + (\mu_\rho + \beta_2)t$

```
modele<-lm(rayon~temps+espece+temps:espece,data=data)
```

ANCOVA et R

- ANCOVA $r_i(t) = (\mu_{r_0} + \alpha_i) + (\mu_\rho + \beta_i)t + \epsilon_i$
- $r_1(t) = (\mu_{r_0} + \alpha_1) + (\mu_\rho + \beta_1)t$
- $r_2(t) = (\mu_{r_0} + \alpha_2) + (\mu_\rho + \beta_2)t$

```
modele<-lm(rayon~temps+espece+temps:espece,data=data)
```

- Coefficients estimés

```
dummy.coef(modele)
```

Full coefficients are

(Intercept):	-0.1192878		
temps:	0.1337302		
espece:		dp	pm
	0.00000000	0.01531703	
temps:espece:		dp	pm
	0.00000000	-0.03526847	

ANCOVA et R

- ANCOVA $r_i(t) = (\mu_{r_0} + \alpha_i) + (\mu_\rho + \beta_i)t + \epsilon_i$
- $r_1(t) = (\mu_{r_0} + \alpha_1) + (\mu_\rho + \beta_1)t$
- $r_2(t) = (\mu_{r_0} + \alpha_2) + (\mu_\rho + \beta_2)t$

```
modele<-lm(rayon~temps+espece+temps:espece,data=data)
```

- Coefficients estimés

```
dummy.coef(modele)
```

Full coefficients are

(Intercept):	-0.1192878		
temps:	0.1337302		
espece:		dp	pm
	0.00000000	0.01531703	
temps:espece:		dp	pm
	0.00000000	-0.03526847	

- $\hat{\mu}_\rho = 0.133$ et $\hat{\mu}_{r_0} = -0.119$
- $\hat{\alpha}_1 = 0$ et $\hat{\alpha}_2 = 0.015$
- $\hat{\beta}_1 = 0$ et $\hat{\beta}_2 = -0.035$

Test des effets globaux

- Quels effets sont significatifs ?

```
anova(modele)
```

Analysis of Variance Table

Response: rayon

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
temps	1	1014.35	1014.35	39405.34	< 2.2e-16 ***
espece	1	53.39	53.39	2073.95	< 2.2e-16 ***
temps:espece	1	23.01	23.01	893.77	< 2.2e-16 ***
Residuals	10816	278.42	0.03		

Variance expliquée par chaque variable

- Quelle est l'influence de chaque variable ?

```
cbind(variable=row.names(anova(modele)),influence=  
anova(modele)$Sum/(sum(anova(modele)$Sum)))
```

	variable	influence
[1,]	"temps"	"0.740854333681361"
[2,]	"espece"	"0.0389919808107421"
[3,]	"temps:espece"	"0.0168035789994021"
[4,]	"Residuals"	"0.203350106508495"

Comparaison des modalités

- Ici on ne s'intéresse qu'à l'effet du génotype sur la pente

Comparaison des modalités

- Ici on ne s'intéresse qu'à l'effet du génotype sur la pente
- Y-a-t'il des différences significatives entre les modalités ?

Comparaison des modalités

- Ici on ne s'intéresse qu'à l'effet du génotype sur la pente
 - Y-a-t'il des différences significatives entre les modalités ?
- comparaison post-hoc des pentes avec `cld()` et `emtrends()`

Comparaison des modalités

- Ici on ne s'intéresse qu'à l'effet du génotype sur la pente
 - Y-a-t'il des différences significatives entre les modalités ?
- comparaison post-hoc des pentes avec `cld()` et `emtrends()`

```
cld(emtrends(modele,"espece",var="temps"))
```

espece	temps.trend	SE	df	lower.CL	upper.CL	.group
pm	0.0985	0.000871	10816	0.0968	0.100	1
dp	0.1337	0.000796	10816	0.1322	0.135	2

Confidence level used: 0.95

significance level used: alpha = 0.05

Présentation des résultats

① Graphiques : modèles ajustés & données

② Tableau d'analyse de la variance : test des effets globaux

Variable	Df	p-value	variance expliquée
Year	3	0.03	6.6 %
Fields	24	< 0.001	3.2 %
Severity	5	0.15	0.5 %

③ Comparaison des modalités (coefs estimés, tests post-hoc, graphiques...)

Trois erreurs courantes

- ❗ Confusion effet sur l'ordonnée à l'origine et sur la pente (i.e. interaction quanti-quali)
- ❗ Test des effets globaux : `Anova()` et non `summary()`

Trois erreurs courantes

- ❶ Confusion effet sur l'ordonnée à l'origine et sur la pente (i.e. interaction quanti-quali)
- ❷ Test des effets globaux : `Anova()` et non `summary()`
- ❸ Hypothèse de croissance radiale constante puis modèle linéaire sur la surface

Trois erreurs courantes

- ❶ Confusion effet sur l'ordonnée à l'origine et sur la pente (i.e. interaction quanti-quali)
- ❷ Test des effets globaux : `Anova()` et non `summary()`
- ❸ Hypothèse de croissance radiale constante puis modèle linéaire sur la surface

Merci de votre attention

Questions ?

melen.leclerc@inrae.fr