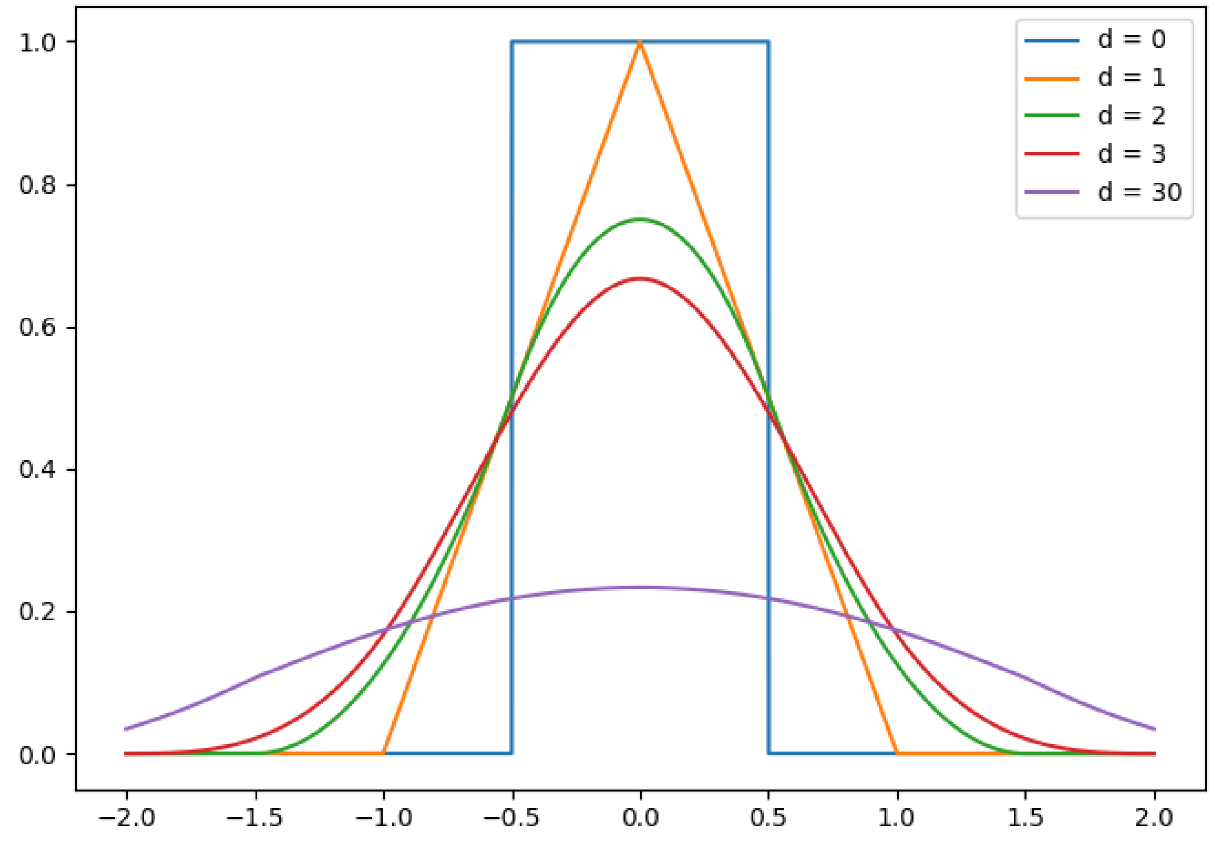
Construction, validation et calibration d’un niveau test statistique

Projet M1



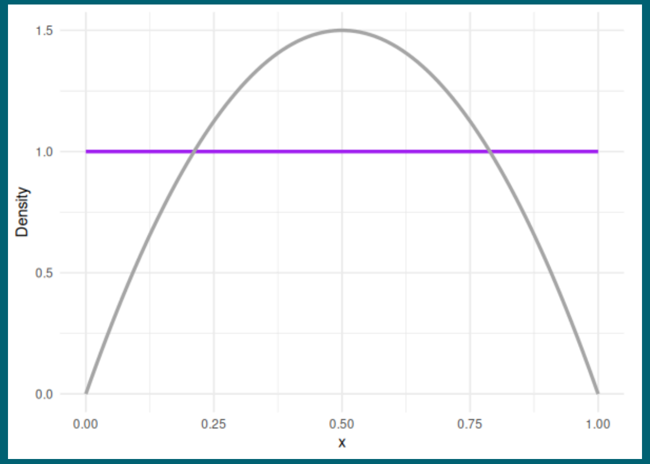
Présenté par :

EL YAMANI MARYEM

STRAUSS DANIEL

Encadré par :

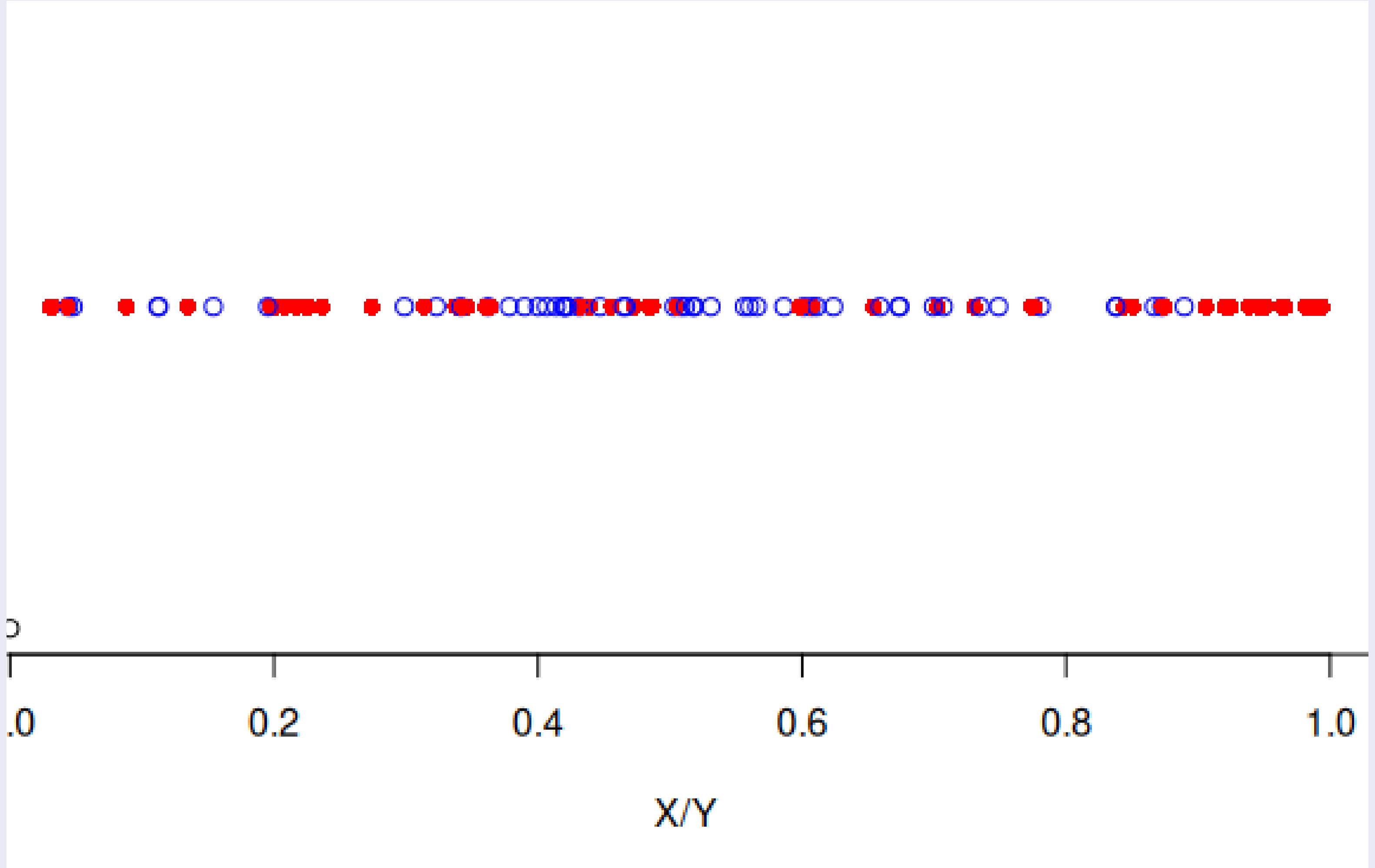
Gilles Ducharme



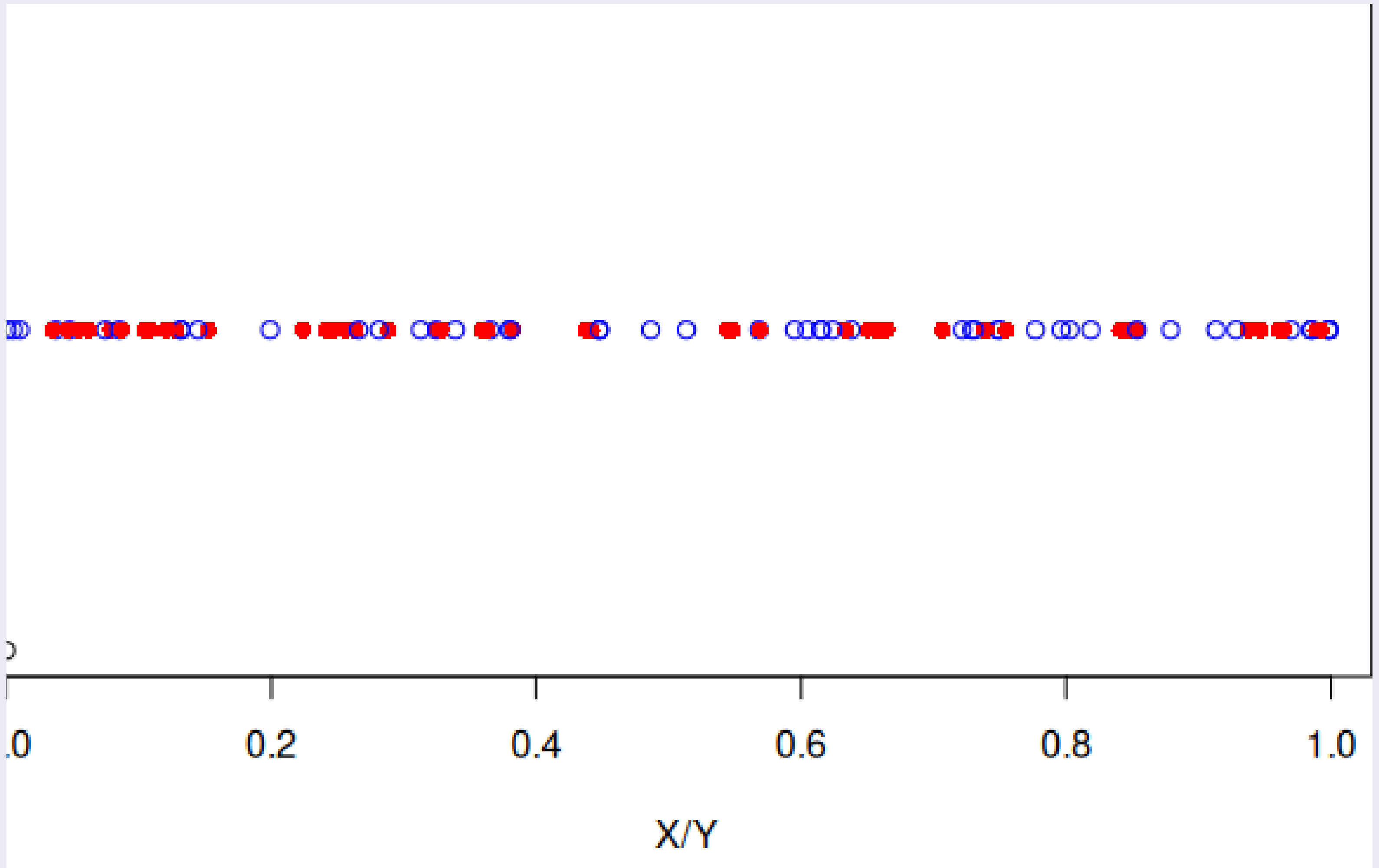
Les densités de Unif(0,1) et de Beta(2,2)

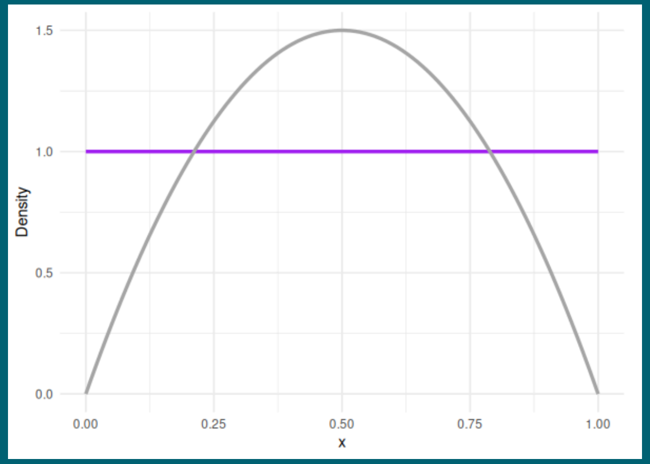
Y_i suit la loi Unif(0,1)
Y_j suit la loi de Beta(2,2)
Vous pouvez deviner à l'aide des graphiques Y_i et Y_j?

X_1 ~ Unif(0,1) | Y_1 ~ ??



X_2 ~ Unif(0,1) | Y_2 ~ ??

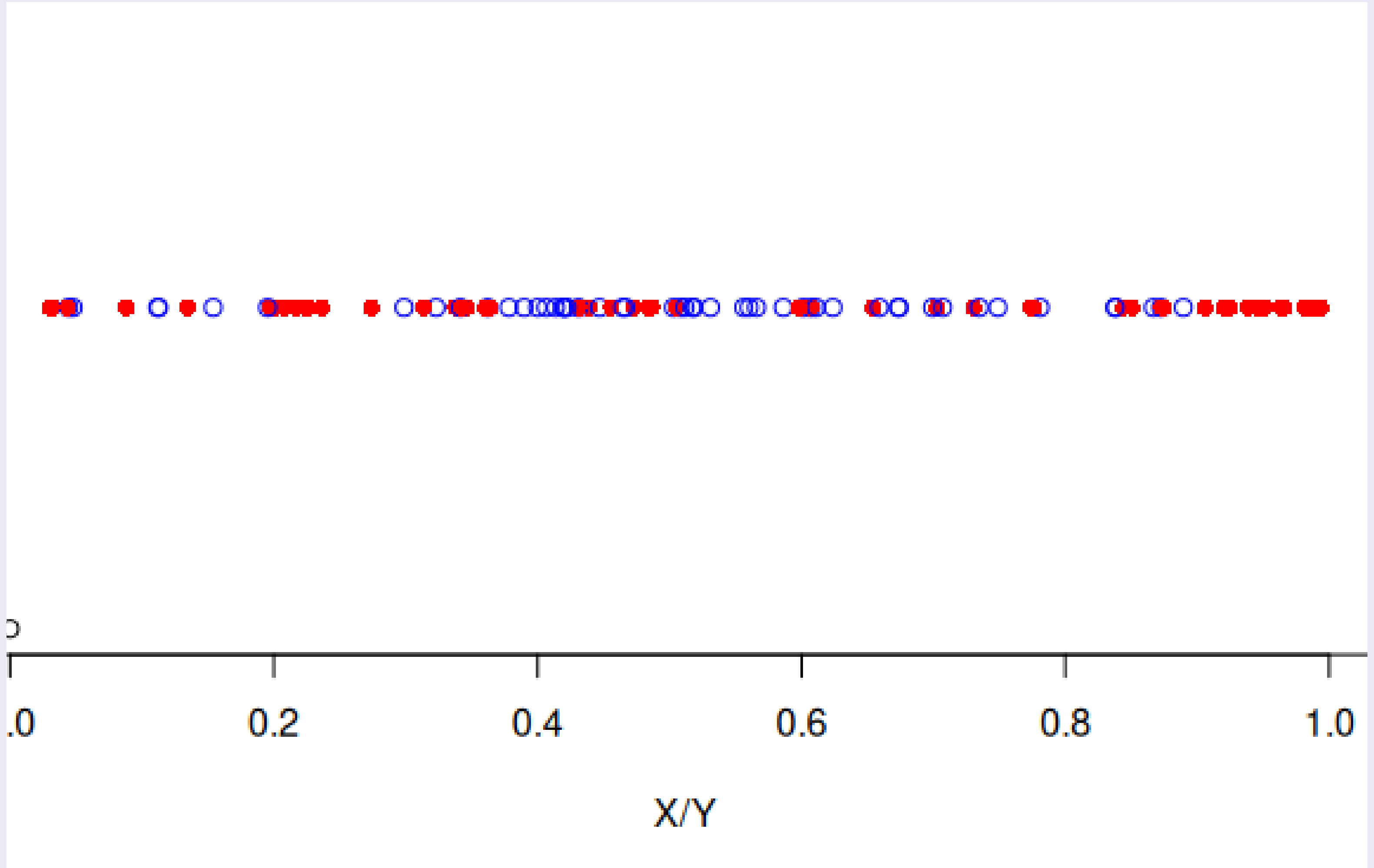




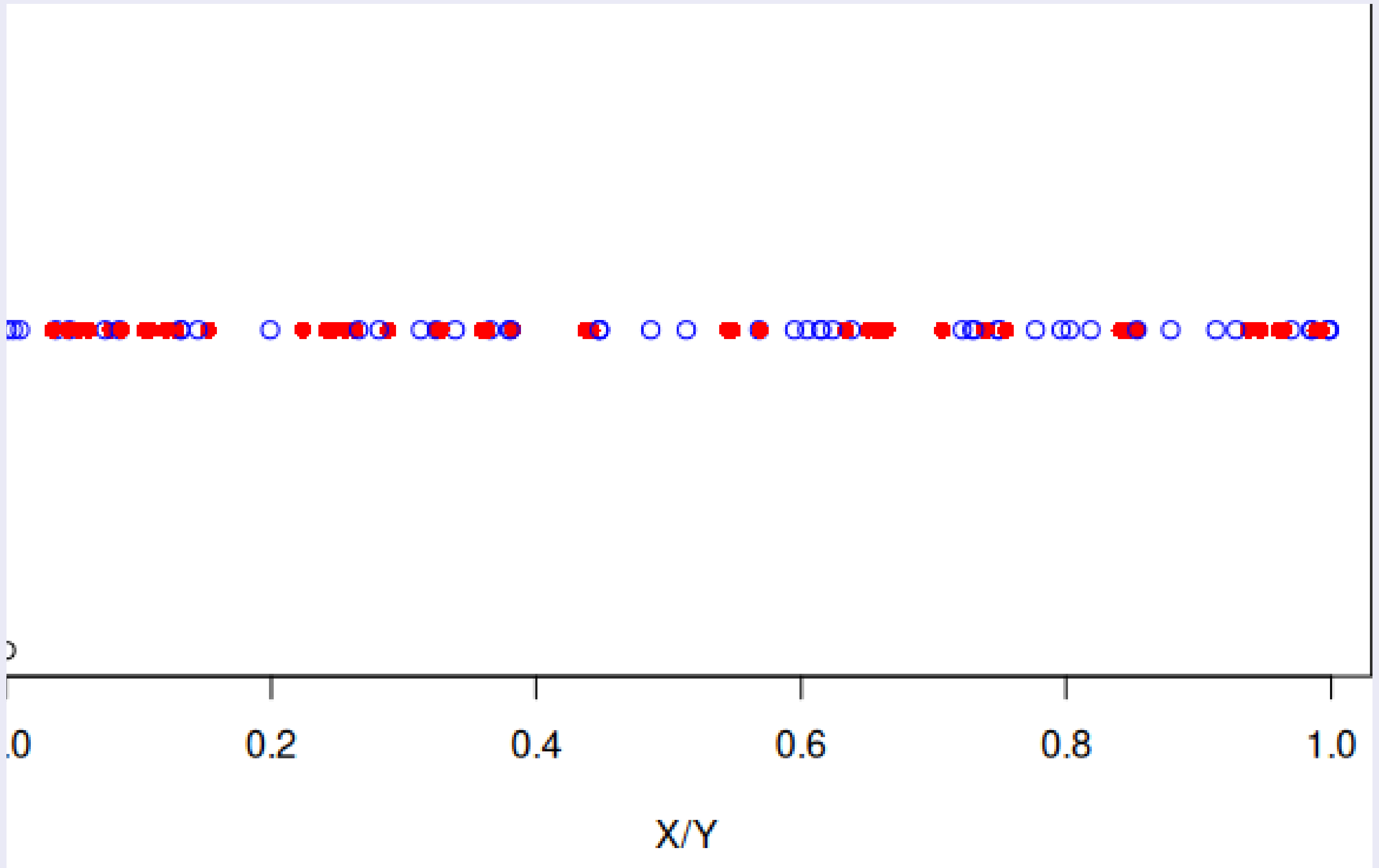
Les densités de Unif(0,1) et de Beta(2,2)

La statistique du test D sera introduite après.
A noté: 95 quantile sous H_0 de D est ~ 30.4

X_1 ~ Unif(0,1) | Y_1 ~ Beta(2,2) | D = 265.5



X_2 ~ Unif(0,1) | Y_2 ~ Unif(0,1) | D = 2.11



Plan adapté

1. Introduction du projet
2. Revue des tests classiques
3. Introduction du nouveau test
4. Validation de notre test Ducharme
5. Conclusion générale

1

Introduction du projet

Notre projet se concentre sur le développement et la validation d'un nouveau test statistique, le test DUCHARME, pour les hypothèses H0 : F = G contre H1 : F ≠ G. Ce test propose une approche novatrice pour évaluer les différences significatives entre les densités de probabilité. Notre objectif principal est de comprendre et d'explorer les étapes cruciales dans le développement de ce test, en examinant en détail les caractéristiques spécifiques qui le distinguent des tests classiques.

2

Revue des tests classiques

Test paramètrique

Test de Student

Ce test est utilisé pour comparer les moyennes de deux échantillons indépendants en supposant qu'ils proviennent de populations normales

Tests non paramètriques

Test de type Cramer-von Mises et notamment d'AndersonDarling

Ces tests statistiques sont utilisés pour évaluer l'ajustement d'un échantillon à une distribution théorique. Ils comparent les fonctions de répartition empirique et théorique pour déterminer si l'échantillon suit la distribution théorique donnée.

Test de Wilcoxon-Mann-Whitney

Ce test est utilisé pour comparer les distributions de deux échantillons indépendants. Il se base sur les rangs des observations pour évaluer la probabilité que les deux échantillons proviennent de la même distribution.

Test de Kolmogorov-Smirnov

Ce test est utilisé pour évaluer l'ajustement d'un échantillon à une distribution théorique. Il compare les fonctions de répartition empirique et théorique pour déterminer si l'échantillon suit la distribution théorique donnée.

Test de Wilcoxon

Hypothèses du test :

$$H_0 : P(X \leq Y) = \frac{1}{2}$$

$$H_1 : \begin{cases} P(X \leq Y) \neq \frac{1}{2} \\ P(X \leq Y) < \frac{1}{2} \\ P(X \leq Y) > \frac{1}{2} \end{cases}$$

Statistique du test :

La somme des rangs de l'enchantallion de X

$$U = \sum_{i=1}^{n} R_i$$

Test de Kolmogorov-Smirnov

Hypothèses du test :

$$H_0 : F = G$$

$$H_1 : F \neq G$$

Statistique du test :

La différence entre les fonctions de répartitions empiriques des deux distribution

$$D_{n,m} := \sup_x |F_n(x) - G_m(x)|$$

Région critique :

$$R_\alpha = (\sqrt{\frac{\ln(\alpha/2)}{2} \cdot \frac{n+m}{n \cdot m}}, \infty)$$

Test statistique :

Décision/Réalité	H_0 vraie	H_1 vraie
Ne pas rejeter H_0	$1-\alpha$	β
Rejeter H_0	α	$1-\beta$

3

Introduction du nouveau test

Notre test statistique nommé **Ducharme** se base sur des noyaux et suivra une approche similaire au test Kolmogorov-Smirnov.

Supposons :

$(X_1, ..., X_n)$ iid de loi f sur [0,1]
$(Y_1, ..., Y_m)$ iid de loi g sur [0,1]

On a :

$$h(x) = \frac{f(x)}{g(x)}$$

On veut alors estimer les densités de notre test par les fonctions splines.

Splines :

Les fonctions splines sont des fonctions mathématiques utilisées dans l'interpolation Splines pour estimer une fonction continue à partir d'un ensemble de points de données. Elles sont définies par des morceaux de polynômes qui sont connectés de manière continue en utilisant des conditions spécifiques.

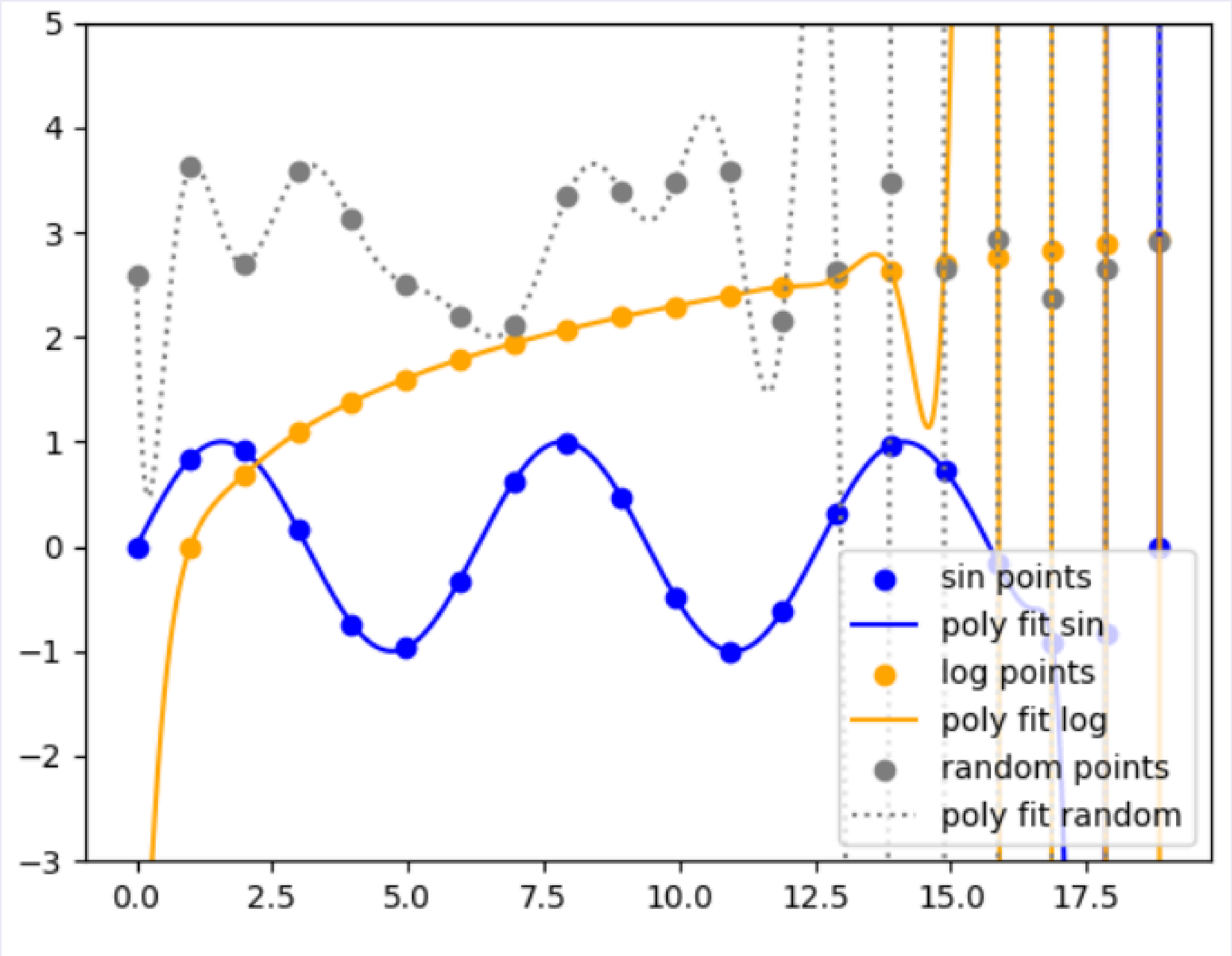


Figure 1 : Ajustement polynomiale

- **Ce graphique montre trois exemples d'ajustement polynomial avec 20 points de données.**
- **On peut remarquer que quand les valeurs de x sont élevés , les fonctions ajustées commencent à mal se comporter**

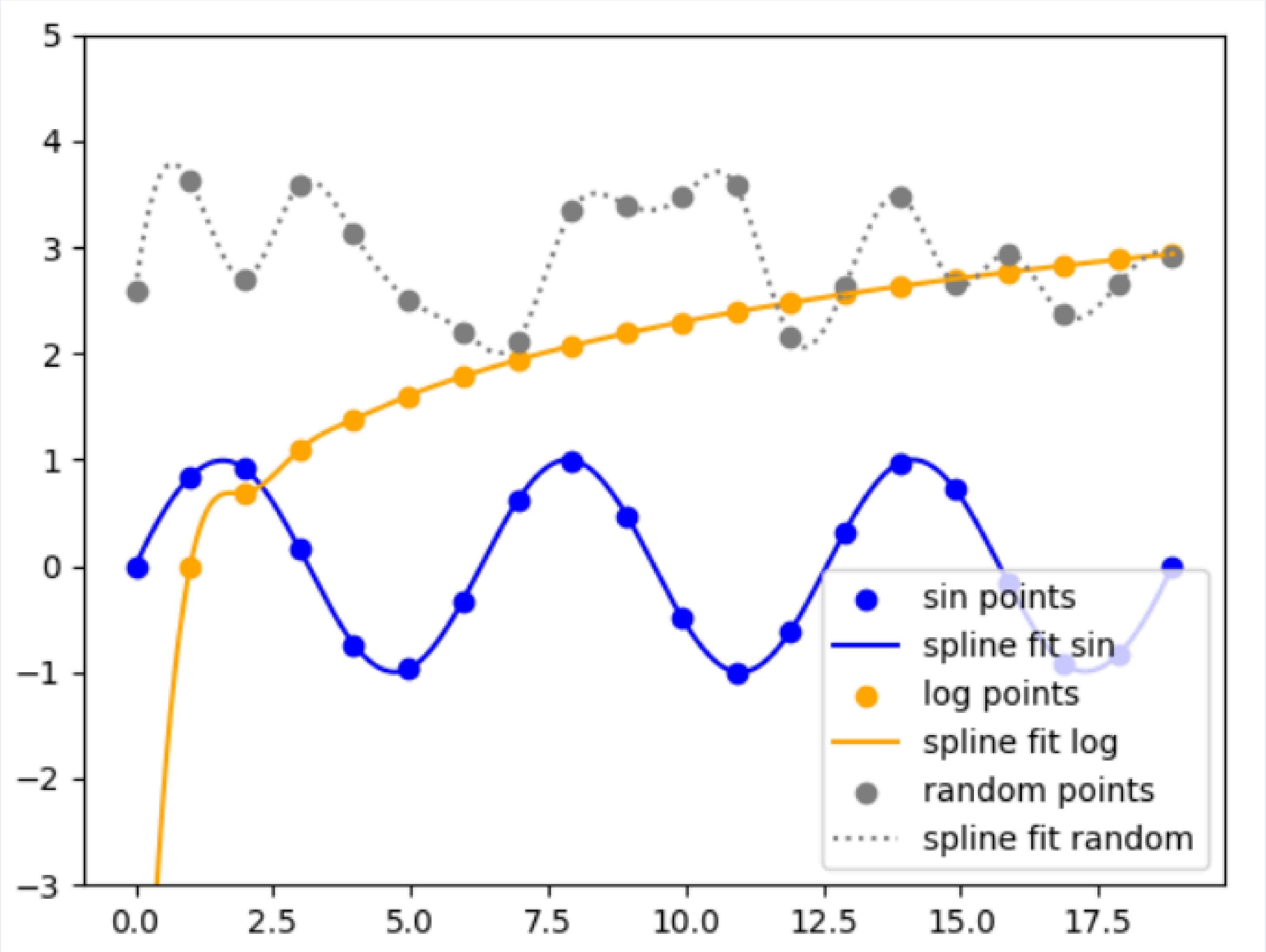


Figure 2: Interpolation splines

- Ce graphique illustre trois exemples d'ajustement par spline avec 20 points de données et degré (d = 3).
- On remarque les splines sont mieux représenté que les ajustements polynomiaux.

Estimation de densité par splines :

On veut estimé la fontion : $h(x) = \frac{f(x)}{g(x)}$

Tout d'abord, nous avons collecté deux échantillons de données :

$$(X_1, ..., X_n) \sim Beta(2,3)$$

$$(Y_1, ..., Y_m) \sim Beta(1,1)$$

D'où :

$$f(x) = Beta(2,3)$$

$$g(x) = Beta(1,1)$$

L'estimateur de h(x) est :

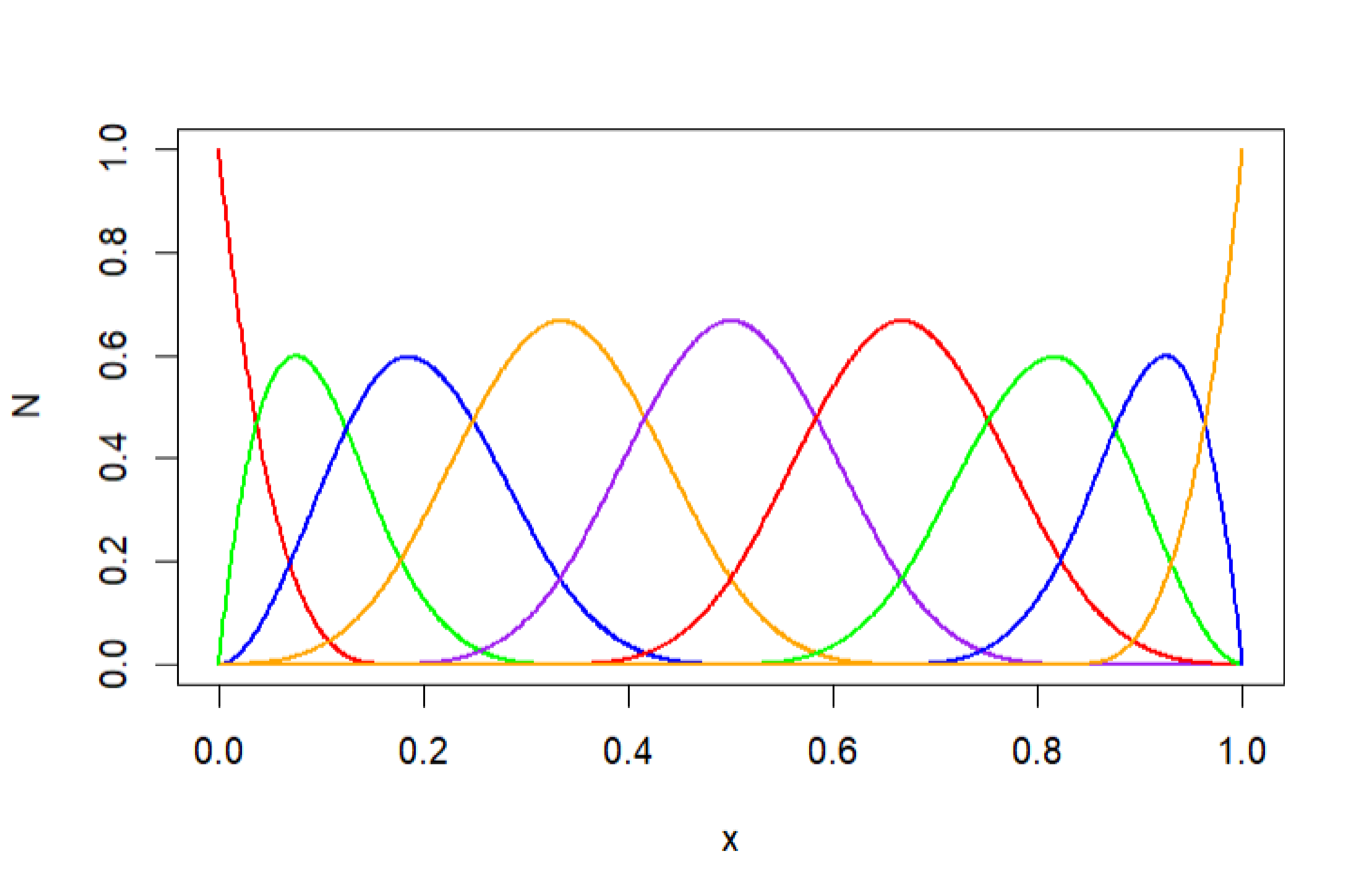
$$\hat{h}(x) = N'(x)\hat{M}_Y^{-1}\frac{1}{n}\sum_{i=1}^{n} N(X_i)$$

Quelques formules :

$$N'(Z_i) := (N_1(Z_i), ..., N_{k+d+1}(Z_i))$$

N(Z_i): Base des B-splines de l'échantillon Z au valeur Z_i

$$\hat{M}_Y^{-1} := \frac{1}{n}\sum_{i=1}^{n} N(Y_i)N'(Y_i)$$



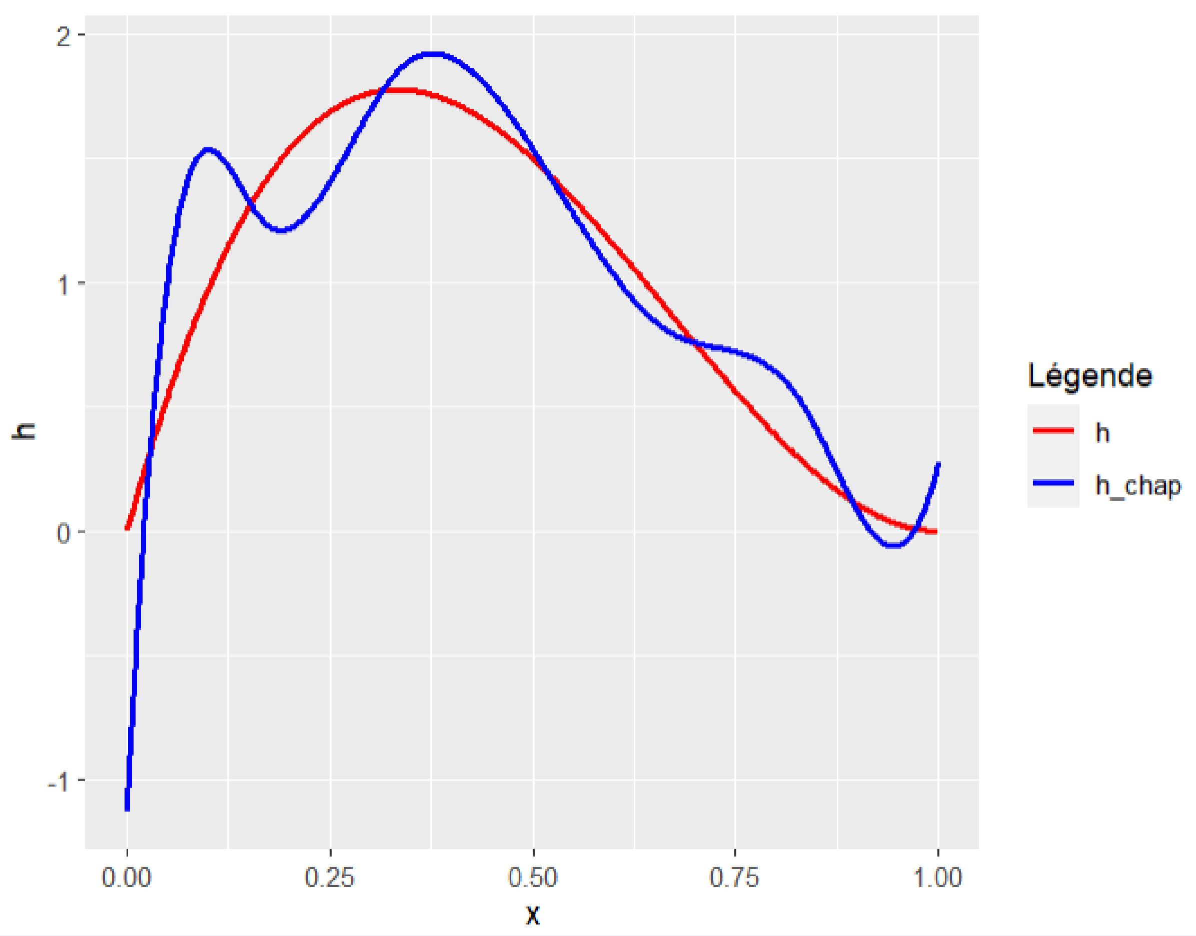


Figure 3 : Graphique des noeuds intérieurs positionnées avec choix de d=3 et k=5

Figure 4 :
Estimation de h(x) pour les densités f= Beta(2,3) , g= Beta(1,1) et pour les choix de k=5 ,d=3

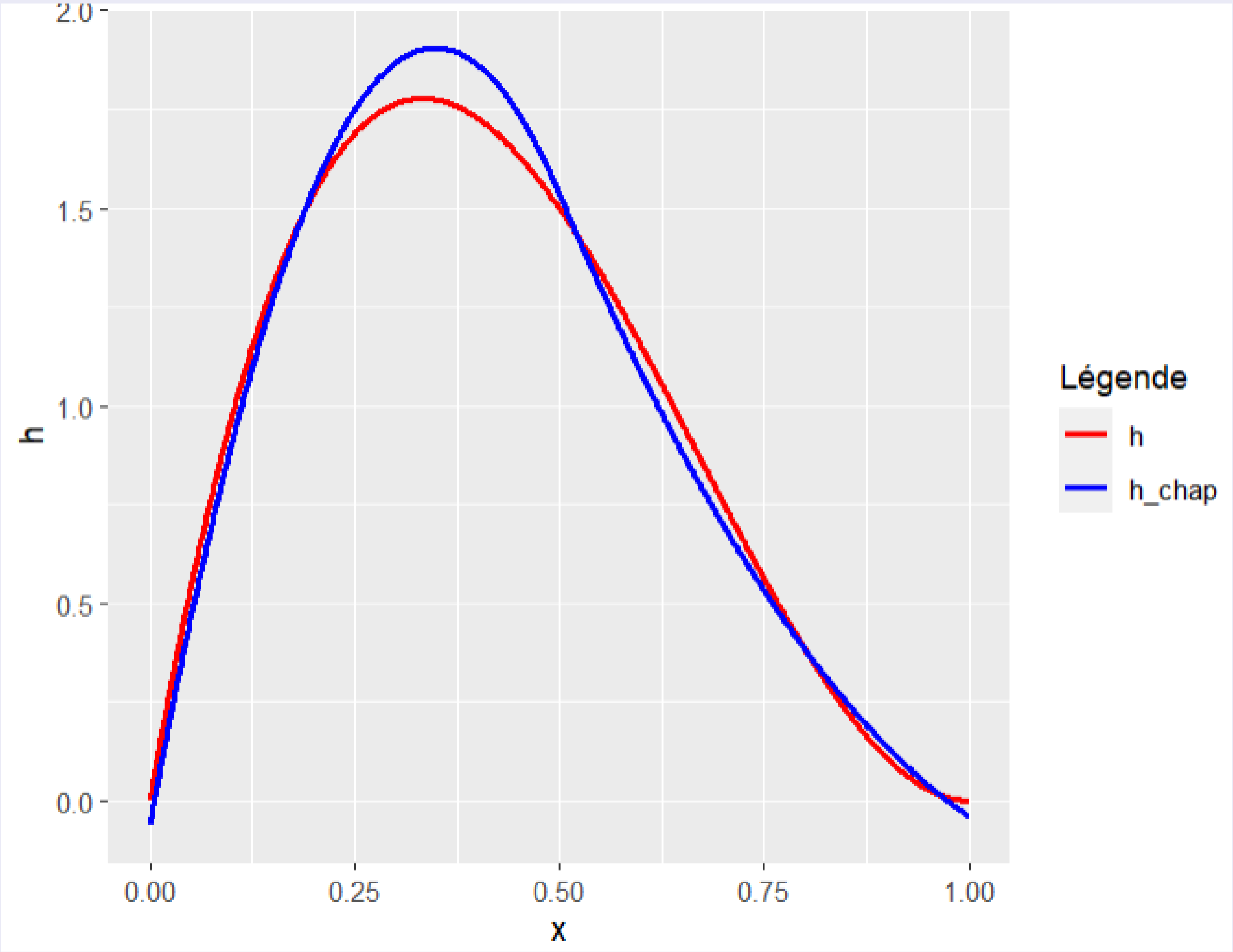


Figure 5 : Estimation des densités pour d = 2 et k = 1

- Cette estimation semble être parfaite.

- L'erreur moyenne quadratique entre les h et son estimation est de :

  0.006177064

Construction du nouveau test :

Hypothèses du test :

$H_0 : f = g$
$H_1 : f \neq g$
Revient donc à tester :
$H_0 : h = 1$
$H_1 : h \neq 1$

Test Ducharme

Statistique du test :

$D = n \int (\hat{h}(x) - 1)^2 dx$
ou encore :
$D = n\mathbf{1}'[\hat{M}_x - \hat{M}_y]\hat{M}_y^{-1} M \hat{M}_y^{-1}[\hat{M}_x - \hat{M}_y]\mathbf{1}$

Nous avons calculé la statistique du test sur RStudio pour les deux échantillons X et Y qui suivent respectivement la loi Beta(1,1)

Région critique :

On va la déterminer dans le prochain chapitre

Déterminer la région critique:

D plus grand => ĥ plus différent de 1=> H_0 a moins de chance d'être valide

=> On choisit la forme suivante pour la région critique: $R_\alpha := [q_\alpha, \infty[$ où q_α est la quantile de $1-\alpha$ de D sous l'hypothèse nulle H_0.

Rappel:

$D = n\int(\hat{h}(x)-1)^2dx$

<u>Comment déterminer q_α ?</u>

On calcule D 10^5 fois, pour X ~ Unif(0,1) et Y~Unif(0,1) et on estime q_α la quantile de $1-\alpha$ pour le résultat de D

4

Validation de notre test Ducharme

Vérifier l'hypothèse nulle H_0 :

Vérifier si dans tous les cas de H_0,
la probabilité de rejeter H_0 est inférieure à alpha.
Si ce n'est pas le cas, ajuster H_0

Validation de la Puissance

Comparez la puissance de notre test à celle d'autres tests similaires
d'autres tests célèbres similaires, comme le Test de Kolmogorov-Smirnovou le Test de Wilcoxon-Mann-Whitney.

Vérifier l'hypothèse nulle H_0 :

Tout d'abord on veut vérifier l'hypothèse nulle H_0 pour X de loi f, Y de loi g:

H_0: f = g

?

X,Y ~ U(0,1): Probabilité de rejet est 0.05

X,Y ~ Beta(1,2): Probabilité de rejet est 0.27

X,Y ~ Beta(2,2):Probabilité de rejet est 0.18

X,Y ~ Beta(1,6): Probabilité de rejet est 0.84

X,Y ~ Beta(6,6): Probabilité de rejet est0.71

=> le seuil critique choisi ne permet pas à notre test d'avoir l'hypothèse zéro de H_0

Nous utiliserons pour l'instant l'hypothèse $\hat{H}_0$ de X ~ Unif(0,1), Y ~ Unif(0,1).
L'extension de notre test pour obtenir H_0 pourrait faire l'objet d'un travail ultérieur.

On a pris les hyperparametres
d = 1
k = 2
(On estime les densités avec les splines de degree d et k kno

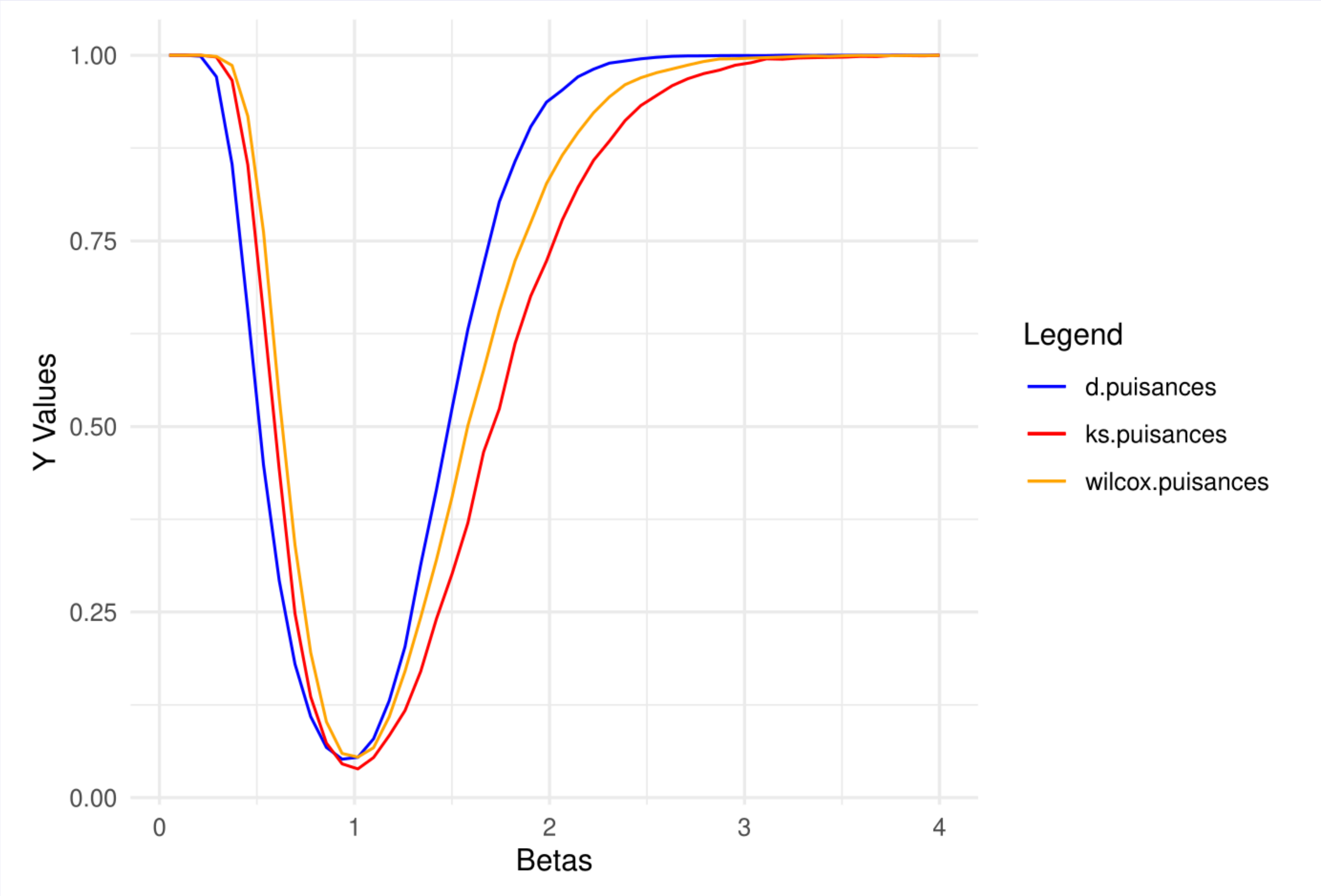
Comparaison de la puissance avec celle des tests classiques :

α	β	D.Puis	Ks.Puis	Wilcox.Puis
0.010	0.010	1.0000	1.0000	0.1145
0.010	1.005	1.0000	1.0000	1.0000
0.010	2.000	1.0000	1.0000	1.0000
1.005	0.010	1.0000	1.0000	1.0000
1.005	1.005	0.0690	0.0365	0.0460
1.005	2.000	0.9483	0.7290	0.8230
2.000	0.010	1.0000	1.0000	1.0000
2.000	1.005	0.9519	0.7200	0.8305
2.000	2.000	0.8105	0.1180	0.0580

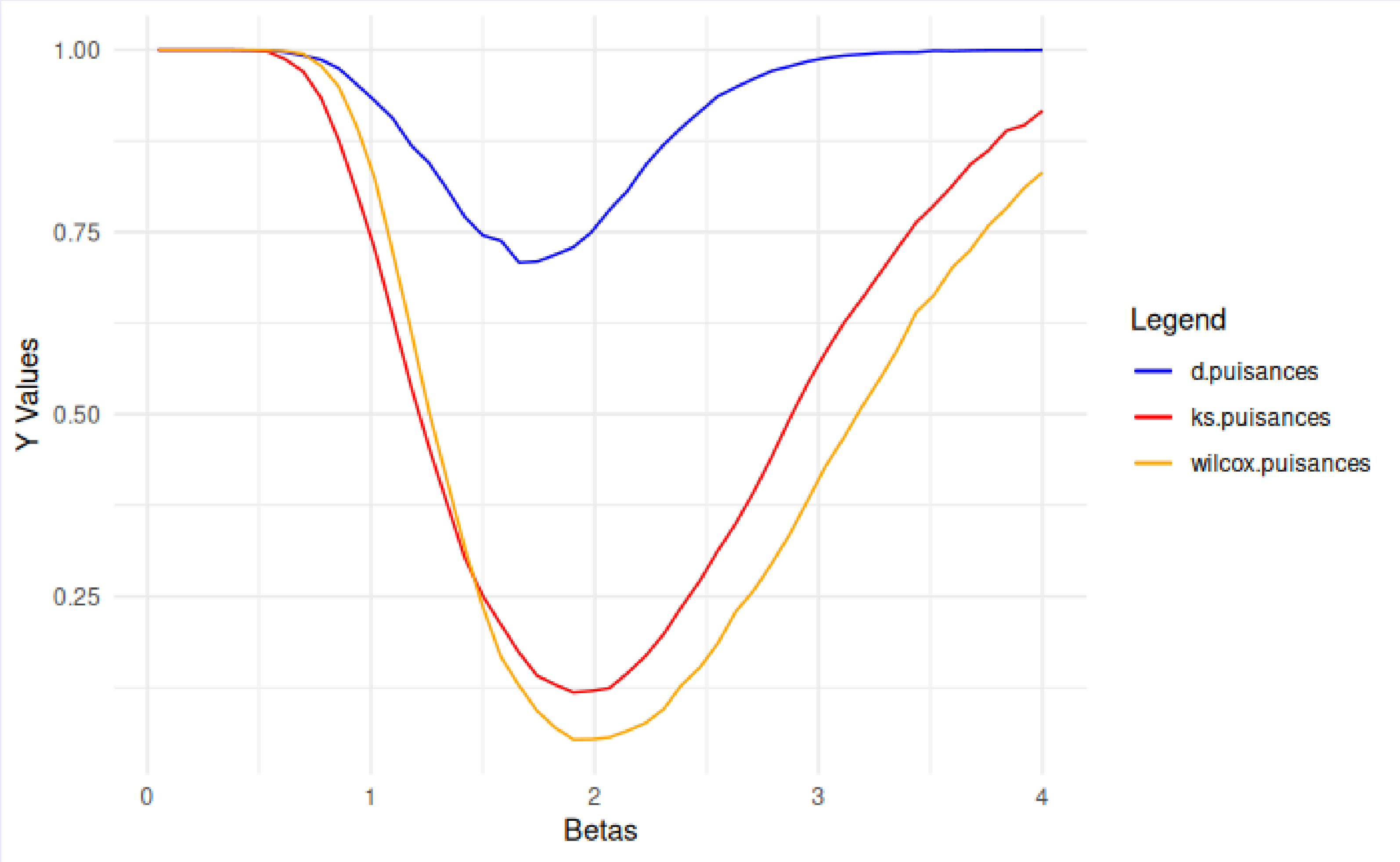
Table 1: Comparaison de la puissance du test de Ducharme avec les puissances du test de Kolmigorov et du test de Wilcoxon pour différents cas
de X ~ Unif(1,1), Y ~ Beta(alpha, beta).

On a pris les hyperparametres
d = 1
k = 2
(On estime les densités avec les splines de degree d et k knots)

Comparaison de la puissance avec celle des tests classiques :

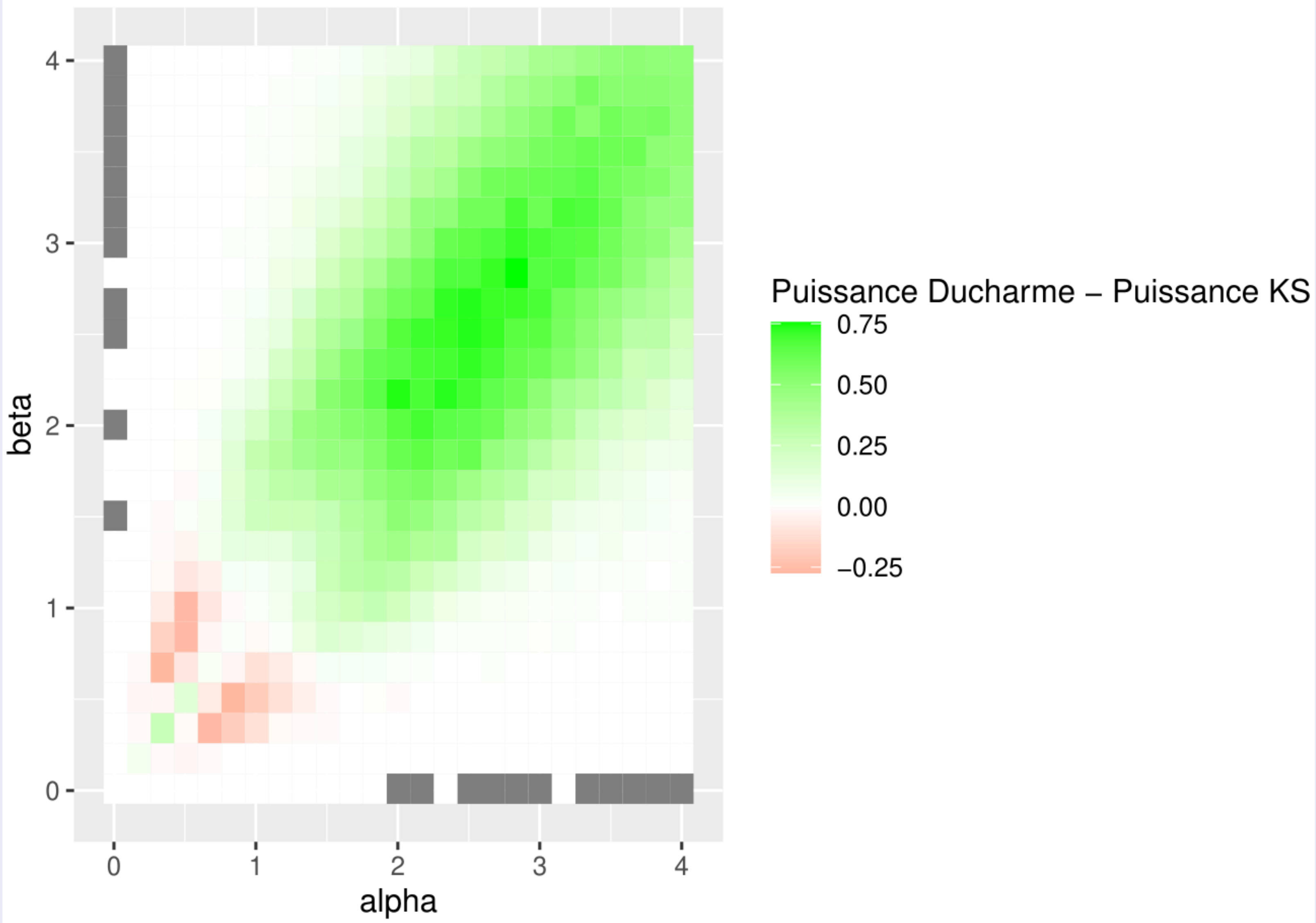


X ~ U(0,1), Y ~ Beta(1, beta)

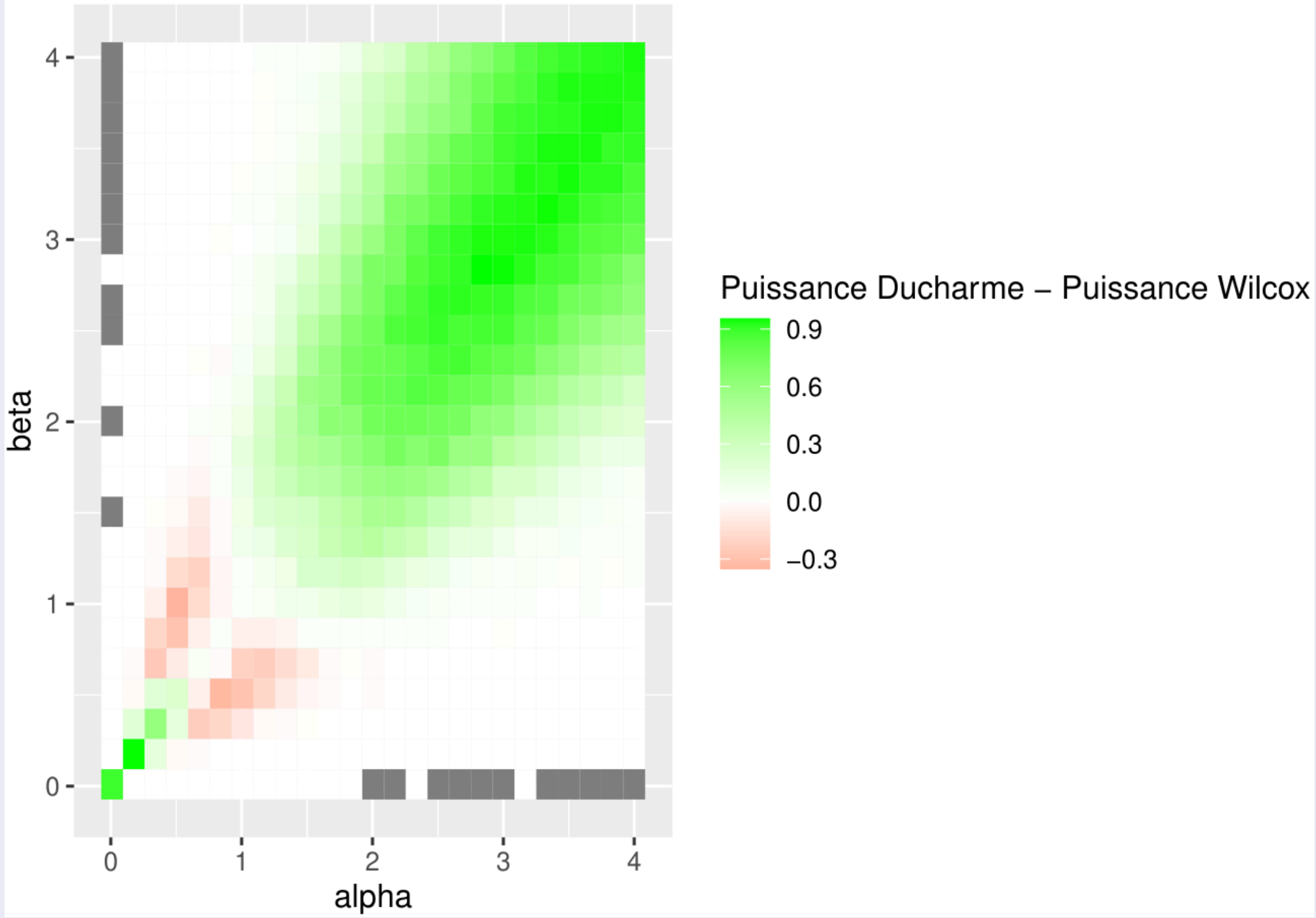


X ~ U(0,1), Y ~ Beta(2, Beta)

Comparer notre test aux autres tests classiques :



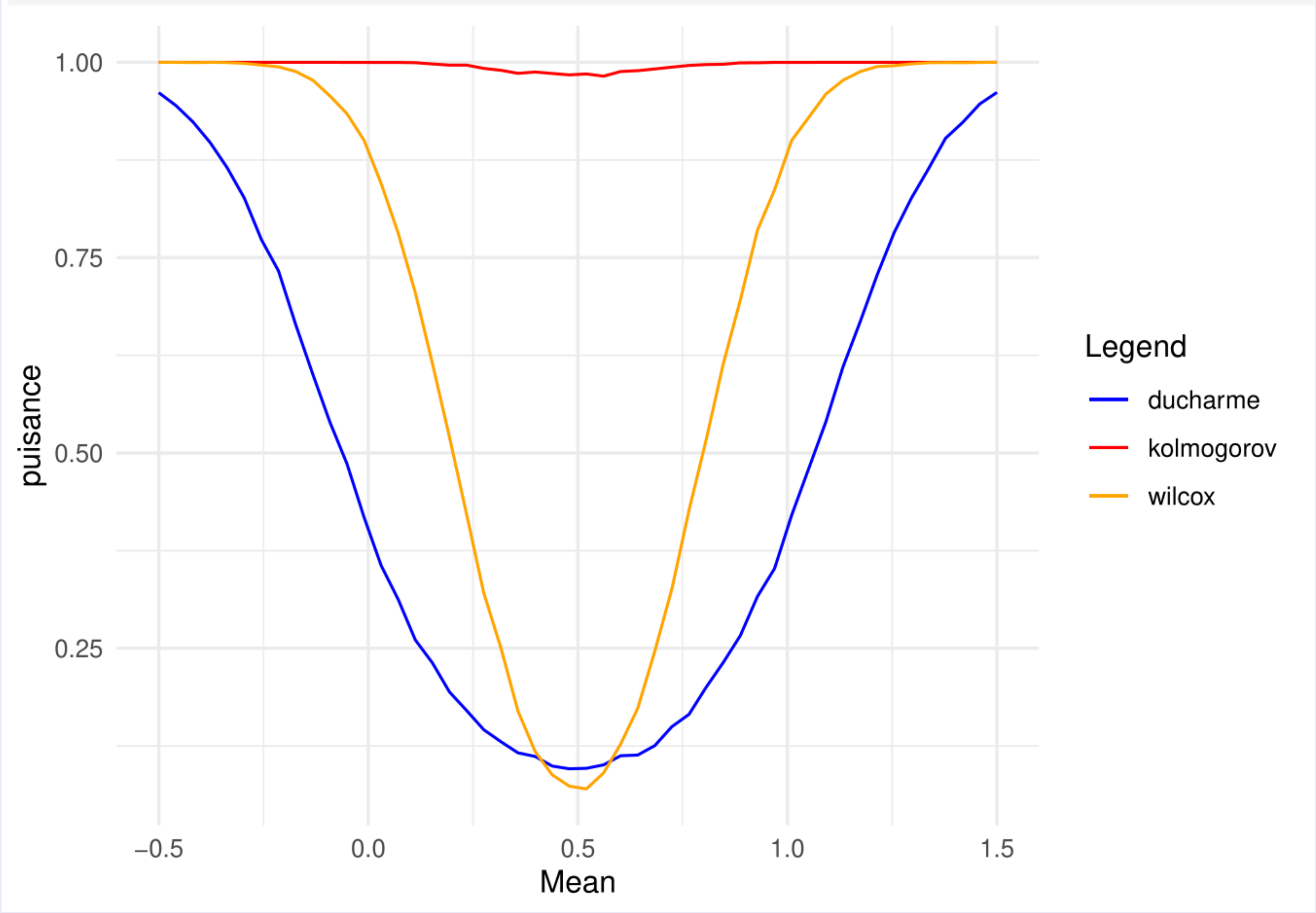
X ~ U(0,1), Y ~ Beta(alpha, beta)



X ~ U(0,1), Y ~ Beta(alpha, beta)

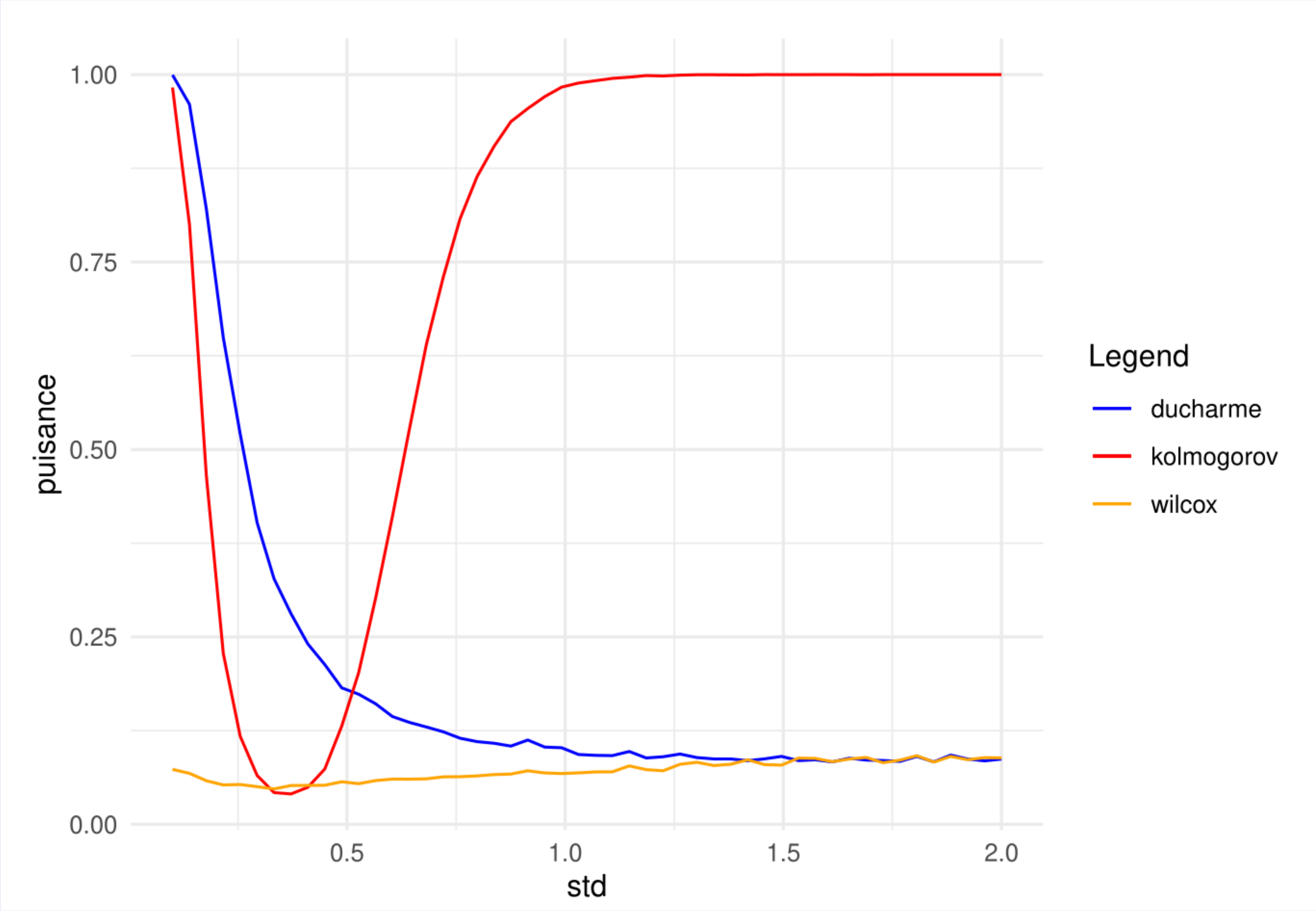
Puissance pour la loi normale

Puissance pour les moyennes differents



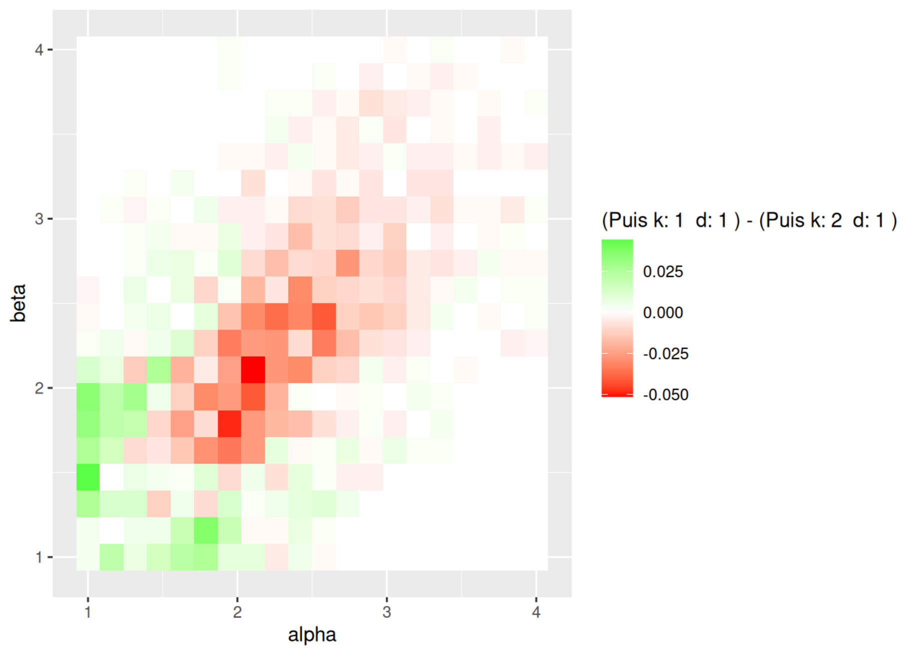
X ~ U(0,1), Y ~ Norm(Mean, 1)

Puissance pour les écarts types differents

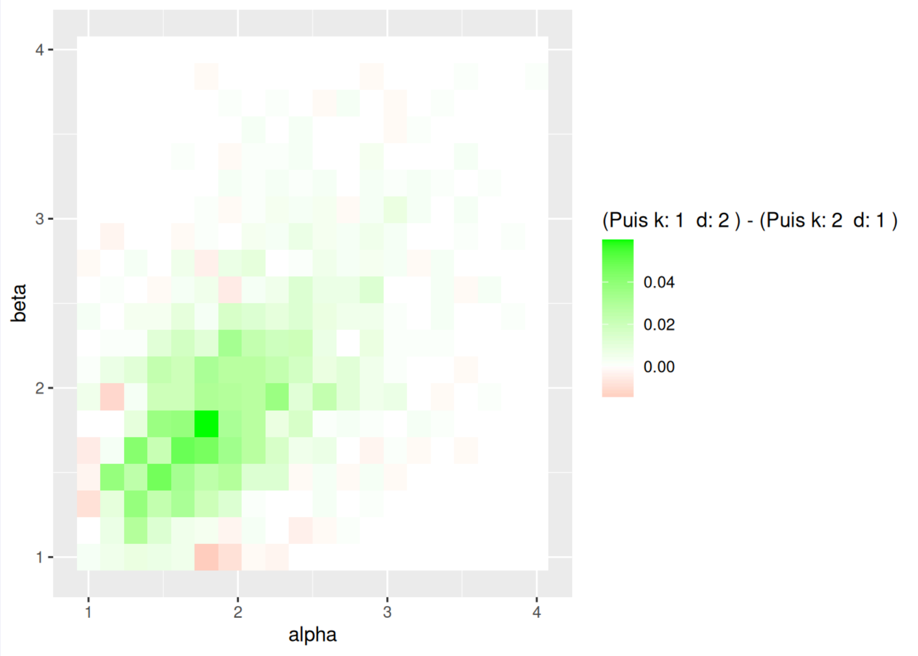


X ~ U(0,1), Y ~ Norm(0,5, std)

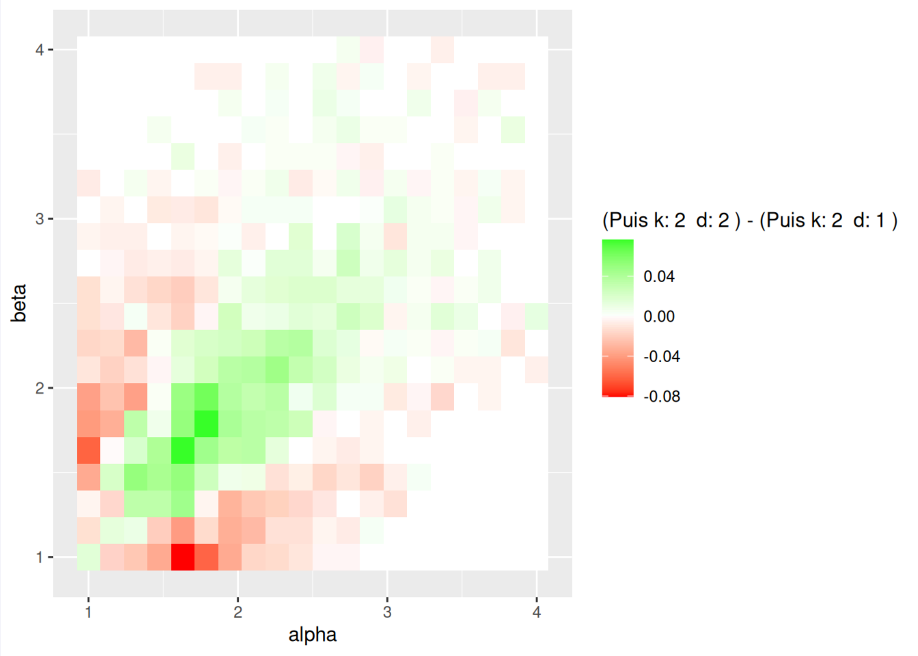
Comparer des puissances pour les differentes valeurs de d et k



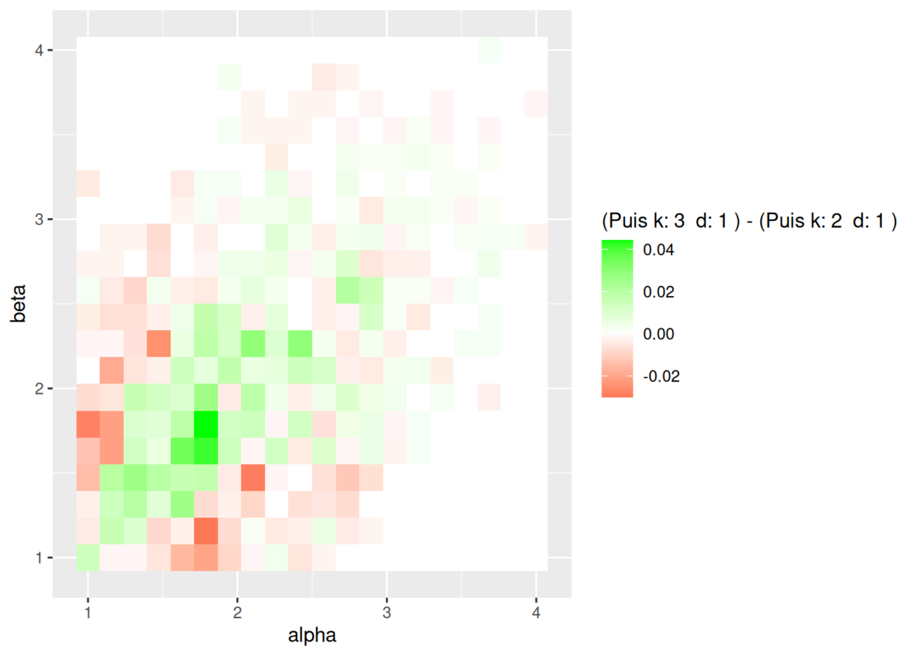
d:1, k: 1



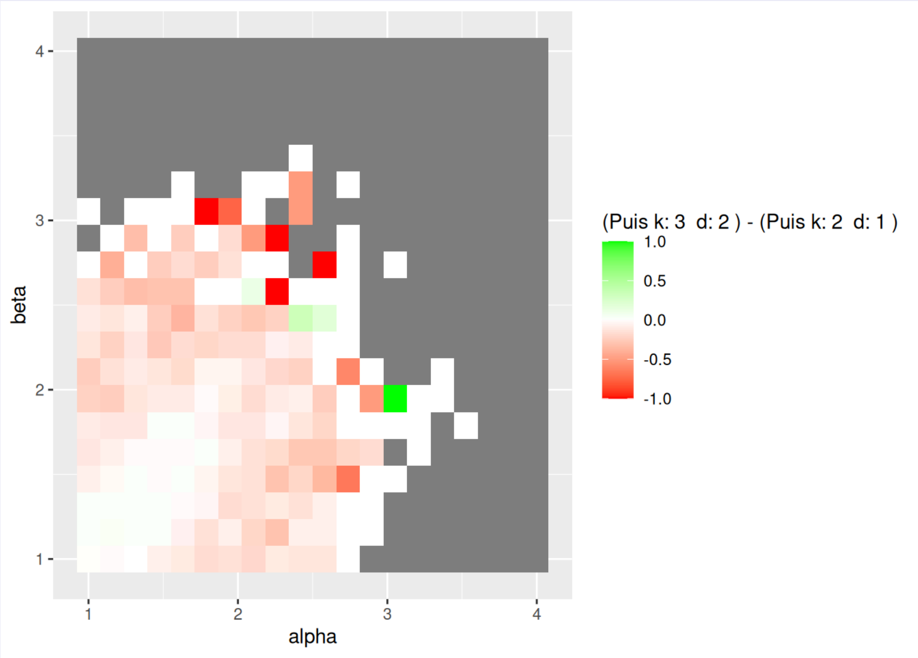
d:2, k: 1



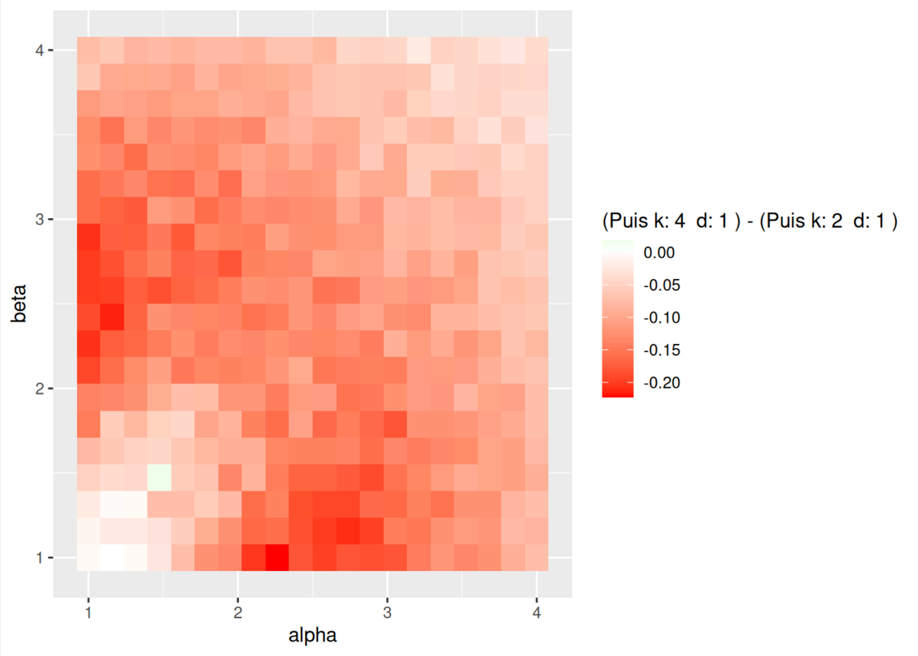
d:2, k: 2



d:1, k: 3



d:2, k: 3



d:1, k: 4

5

Conclusion générale

Merci pour votre attention