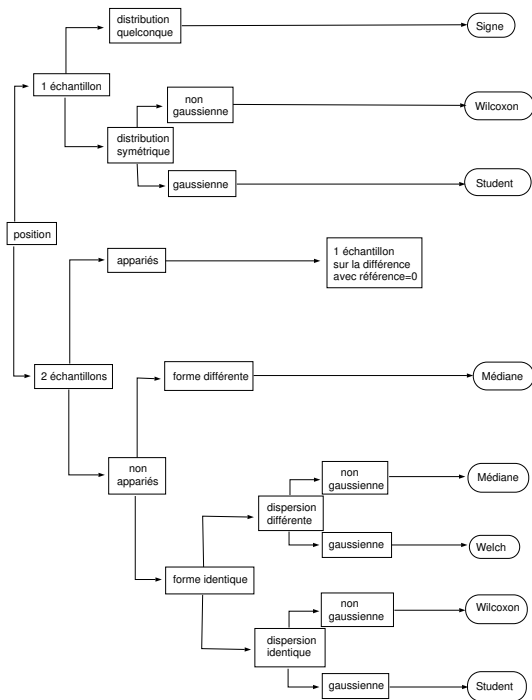


Arbre de NESI



FPSTAT 2 – La décision statistique. 4. Exemples d'applications de tests. 1

Institut National de la Recherche Agronomique

PROBLÈME : Comparaison des hauteurs des arbres deux types de forêts.

Dans deux types de forêts distincts, on a mesuré en mètre les hauteurs dominantes respectivement de 13 et 14 peuplements du même âge choisis au hasard et indépendamment. Les deux forêts sont-elles identiques ?

Type 1	Type 2
23.4	22.5
24.4	22.9
24.6	23.7
24.9	24.0
25.0	24.4
26.2	24.5
26.3	25.3
26.8	26.0
26.8	26.2
26.9	26.4
27.0	26.7
27.6	26.9
27.7	27.4
	28.5

FPSTAT 2 – La décision statistique. 4. Exemples d'applications de tests. 2

Institut National de la Recherche Agronomique

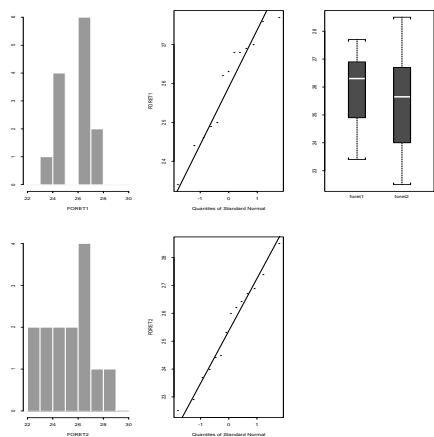


Figure 1: Comparaison des hauteurs des arbres deux types de forêts.

Comparaison des hauteurs des arbres

Sortie Splus (test de Fisher et test de Student)

Test de Fisher : F test for variance equality

```
data: FORET1 and FORET2
F = 0.587, num df = 12, denom df = 13, p-value = 0.3647
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.1861495 1.9013248
sample estimates:
variance of x variance of y
 1.845641      3.144396
```

Test bilatéral de Student : Standard Two-Sample t-Test

```
data: FORET1 and FORET2
t = 0.9542, df = 25, p-value = 0.3491
alternative hypothesis: true difference in means is not equal to 0
95 percent confidence interval:
-0.6759945 1.8430275
sample estimates:
mean of x mean of y
 25.96923 25.38571
```

Comparaison des hauteurs des arbres

Programme Splus (exploration graphique et tests)

```
# EXPLORATION GRAPHIQUE

FORET1_c(23.4,24.4,24.6,24.9,25.0,26.2,26.3,26.8,26.8,26.9,27.0,27.6,27.7)
FORET2_c(22.5,22.9,23.7,24.0,24.4,24.5,25.3,26.0,26.2,26.4,26.7,26.9,
27.4,28.5)

par(mfcol=c(2,2))
hist(FORET1,breaks=22:30)
hist(FORET2,breaks=22:30)
boxplot(FORET1,FORET2,names=c("foret1","foret2"))
#TEST DE MEME DISTRIBUTION et NORMALE
foret.qqnorm(FORET1,FORET2)

# TEST DE FISHER : egalite de variance
foret.vtest_var.test(FORET1,FORET2)
print(foret.vtest)

# TEST DE STUDENT
# test bilatéral de 2 échantillons de forêt non appariés
# test de l'égalité de 2 moyennes dans le cas de variances égales
foret.test_t.test(FORET1,FORET2)
print(foret.test)
```

FPSTAT 2 – La décision statistique. 4. Exemples d'applications de tests. 5

Institut National de la Recherche Agronomique

Comparaison des hauteurs des arbres

Sortie SAS (test de Student)

TTEST PROCEDURE

Variable: HAUT

TYPE	N	Mean	Std Dev	Std Error
1	13	25.96923077	1.35854372	0.37679223
2	14	25.38571429	1.77324437	0.47391949

Variiances	T	DF	Prob> T
Unequal	0.9638	24.2	0.3447
Equal	0.9542	25.0	0.3491

For H0: Variances are equal, F' = 1.70 DF = (13,12) Prob>F' = 0.3647

FPSTAT 2 – La décision statistique. 4. Exemples d'applications de tests. 6

Institut National de la Recherche Agronomique

Comparaison des hauteurs des arbres

Programme SAS (test de Student)

```
data hauteur;
input haut type;
cards;
23.4 1 /* la hauteur est triee par ordre croissant suivant le type */
24.4 1 proc sort data=hauteur;
24.6 1 by type;
24.9 1 run;
25.0 1 /* test de student */
26.2 1 proc ttest data=hauteur;
26.3 1 class type;
26.8 1 var haut;
26.8 1 run;
26.9 1
27.0 1
27.6 1
27.7 1
22.5 2
22.9 2
23.7 2
24.0 2
24.4 2
24.5 2
25.3 2
26.0 2
26.2 2
26.4 2
26.7 2
26.9 2
27.4 2
28.5 2
;
```

PROBLÈME : Comparaison du taux horaire de diminution de sucres sanguins chez des lapins.

Carter et Hubert (1987) présentent des données sur le taux horaire de diminution de sucres sanguins dans deux groupes (lots) de lapins soumis à deux doses d'un médicament. Peut-on prouver une différence de réponse ?

Dose A	0.21	-16.20	-10.10	-9.67	-11.13	1.96	-10.19	-15.87	-12.81
Dose B	1.59	2.66	-6.27	-2.32	-10.87	7.23	-3.76	3.02	15.01

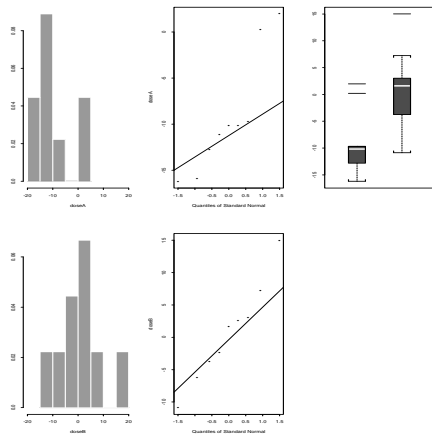


Figure 2: Comparaison du taux horaire de diminution de sucres sanguins chez des lapins.

Comparaison du taux horaire de diminution de sucres sanguins chez des lapins

Sortie Spls (test de somme des rangs de Wilcoxon équivalent au test de Mann-Withney)

Exact Wilcoxon rank-sum test

```
data: doseA and doseB
rank-sum statistic W = 57, n = 9, m = 9, p-value = 0.0106
alternative hypothesis: true mu is not equal to 0
```

Comparaison du taux horaire de diminution de sucres sanguins chez des lapins

Programme Splus (exploration graphique et tests)

```
# -----  
# PROGRAMME : sucre.pg  
# -----  
  
# Exploration Graphique  
  
doseA_c(0.21,-16.2,-10.1,-9.67,-11.13,1.96,-10.19,-15.87,-12.81)  
doseB_c(1.59,2.66,-6.27,-2.32,-10.87,7.23,-3.76,3.02,15.01)  
  
par(mfcol=c(2,3))  
hist(doseA,xlim=c(-20,20),prob=T)  
hist(doseB,xlim=c(-20,20),prob=T)  
  
sucre.qtest_qnorm(doseA)  
qqline(doseA)  
sucre.qtest_qnorm(doseB)  
qqline(doseB)  
boxplot(doseA,doseB)  
  
# TEST DE WILCOXON  
# equivalent au test de mann-Whitney  
  
sucre.wilcox_wilcox.test(doseA,doseB,alternative="two.sided",paired=F)  
print(sucre.wilcox)
```

Comparaison du taux horaire de diminution de sucres sanguins chez des lapins

Sortie SAS (test de somme des rangs de Wilcoxon équivalent au test de Mann-Withney)

```
      N P A R I W A Y   P R O C E D U R E  
  
      Wilcoxon Scores (Rank Sums) for Variable TAUX  
      Classified by Variable DOSE  
  
      DOSE          N      Sum of      Expected      Std Dev      Mean  
      DOSE          N      Scores      Under H0      Under H0      Score  
  
      A             9         57.0      85.5000000    11.3247517    6.3333333  
      B             9        114.0      85.5000000    11.3247517    12.6666667  
  
      Wilcoxon 2-Sample Test (Normal Approximation)  
      (with Continuity Correction of .5)  
  
      S= 57.0000    Z= -2.47246    Prob > |Z| =    0.0134  
  
      T-Test approx. Significance =    0.0243  
  
      Kruskal-Wallis Test (Chi-Square Approximation)  
      CHISQ= 6.3333    DF= 1    Prob > CHISQ=    0.0118
```

Comparaison du taux horaire de diminution de sucres sanguins chez des lapins

Programme SAS (tests)

```
data sucre;
input taux dose $;
cards;
  0.21 A
-16.20 A /* test de Mann-Whitney equivalent a Wilcoxon somme des rangs */
-10.10 A   proc npariway data=sucre;
-9.67 A   class dose;
-11.13 A   var taux;
  1.96 A   run;
-10.19 A
-15.87 A
-12.81 A
  1.59 B
  2.66 B
 -6.27 B
 -2.32 B
-10.87 B
  7.23 B
 -3.76 B
  3.02 B
 15.01 B
;
```

PROBLÈME : Changement d'isolation

La firme Hotpot Stoves a une méthode standard pour isoler ses fours. Pour tester son efficacité, on prend des échantillons au hasard dans la chaîne de production, les fours sont mis en route jusqu'à atteindre 400°C , et on note les temps (en minutes) mis pour le refroidissement à 350°C après avoir coupé le courant sur un échantillon aléatoire de 8 fours. On décide d'étudier une forme moins coûteuse d'isolation, et on l'emploie sur un autre échantillon de 9 fours. On note les temps nécessaires à la même baisse de température sur cette nouvelle isolation. La firme a-t-elle raison de dire qu'il n'y a pas de preuve solide d'une différence de perte de chaleur ?

Isolation Standard	Nouvelle Isolation
15,7	13,7
14,8	14,1
14,2	14,7
16,1	15,4
15,3	15,6
13,9	14,4
17,2	12,9
14,9	15,1
	14,0

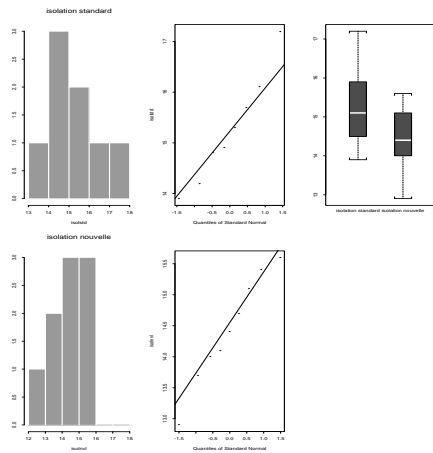


Figure 3: Comparaison de deux méthodes d'isolation.

Comparaison de deux méthodes d'isolation

Sortie Splus (tests)

F test for variance equality

```
data: isolstd and isolnvl
F = 1.5198, num df = 7, denom df = 8, p-value = 0.5685
alternative hypothesis: true ratio of variances is not equal to 1
95 percent confidence interval:
 0.3355948 7.4458313
sample estimates:
variance of x variance of y
 1.139821      0.75
```

Exact Wilcoxon rank-sum test

```
data: isolstd and isolnvl
rank-sum statistic W = 88, n = 8, m = 9, p-value = 0.0694
alternative hypothesis: true mu is greater than 0
```

Standard Two-Sample t-Test

```
data: isolstd and isolnvl
t = 1.7676, df = 15, p-value = 0.0487
alternative hypothesis: true difference in means is greater than 0
95 percent confidence interval:
 0.006846117      NA
sample estimates:
mean of x mean of y
 15.2625 14.43333
```


Comparaison de deux méthodes d'isolation

Programme Splus (exploration graphique et tests)

```
# EXPLORATION GRAPHIQUE
isolstd_c(15.7,14.8,14.2,16.1,15.3,13.9,17.2,14.9)
isolnvl_c(13.7,14.1,14.7,15.4,15.6,14.4,12.9,15.1,14.0)

par(mfcol=c(2,3))
hist(isolstd,breaks=13:18,main="isolation standard")
hist(isolnvl,breaks=12:18,main="isolation nouvelle")
isolstd.qtest.qnorm(isolstd)
qqline(isolstd)
isolnvl.qtest.qnorm(isolnvl)
qqline(isolnvl)
boxplot(isolstd,isolnvl,names=c("isolation standard","isolation nouvelle"))

# TEST DE FISHER : egalite de variance
four.vtest_var.test(isolstd,isolnvl)
print(four.vtest)

# TEST DE STUDENT
# test unilatéral de 2 échantillons de fours non appariés
# dans le cas de variance égales
four.test_t.test(isolstd,isolnvl,alternative="greater",paired=F)
print(four.test)

# TEST DE MANN-WITHNEY-WILCOXON
four.wtest_wilcox.test(isolstd,isolnvl,alternative="greater",paired=F)
print(four.wtest)
```

Comparaison de deux méthodes d'isolation

Sortie SAS (tests)

NPARIWAY PROCEDURE

Wilcoxon Scores (Rank Sums) for Variable TEMPS
Classified by Variable TYPE

TYPE	N	Sum of Scores	Expected Under H0	Std Dev Under H0	Mean Score
1	8	88.0	72.0	10.3923048	11.0000000
2	9	65.0	81.0	10.3923048	7.2222222

Wilcoxon 2-Sample Test (Normal Approximation)
(with Continuity Correction of .5)

S= 88.0000 Z= 1.49149 Prob > |Z| = 0.1358

T-Test approx. Significance = 0.1553

Kruskal-Wallis Test (Chi-Square Approximation)
CHISQ= 2.3704 DF= 1 Prob > CHISQ= 0.1237

N P A R I W A Y P R O C E D U R E

Kolmogorov-Smirnov Test for Variable TEMPS
Classified by Variable TYPE

TYPE	N	EDF at maximum	Deviation from Mean at maximum
1	8	0.25000000	-.623917748
2	9	0.66666667	0.588235294
-----	-----	-----	
	17	0.470588235	

Maximum Deviation occurred at Observation 11
Value of TEMPS at maximum 14.7000000

Kolmogorov-Smirnov 2-Sample Test (Asymptotic)

KS = 0.207973 D = 0.416667
KSa = 0.857493 Prob > KSa = 0.4540

N P A R I W A Y P R O C E D U R E

Cramer-von Mises Test for Variable TEMPS
Classified by Variable TYPE

TYPE	N	Summed Deviation from Mean
1	8	0.134948097
2	9	0.119953864

Cramer-von Mises Statistic (Asymptotic)

CM = 0.014994 CMa = 0.254902

TTEST PROCEDURE

Variable: TEMPS

TYPE	N	Mean	Std Dev	Std Error
1	8	15.26250000	1.06762420	0.37746216
2	9	14.43333333	0.86602540	0.28867513

Variiances	T	DF	Prob> T
Unequal	1.7449	13.5	0.1037
Equal	1.7676	15.0	0.0975

For H0: Variances are equal, F' = 1.52 DF = (7,8) Prob>F' = 0.5685

Comparaison de deux méthodes d'isolation

Programme SAS (tests)

```
data four;
input temps type;
cards;
15.7 1
14.8 1
14.2 1      /* tests Wilcoxon, Smirnov, Kolmogorov */
16.1 1      proc nparlway;
15.3 1      class type;
13.9 1      var temps;
17.2 1      run;
14.9 1
13.7 2
14.1 2      /* test de Student */
14.7 2      proc ttest data=four;
15.4 2      class type;
15.6 2      var temps;
14.4 2      run;
12.9 2
15.1 2
14.0 2
;
```

PROBLÈME : Examen

Lors d'un examen INRA, des candidats sont convoqués par un jury.
On veut tester l'hypothèse selon laquelle la probabilité d'être reçu ne dépend pas du jury.

	Jury 1	Jury 2	Jury 3
Nombre de reçus	50	47	56
Nombre de refusés	5	14	8

Examen

Sortie et Programme Splus (test du Khi2)

Pearson's chi-square test without Yates'
continuity correction

```
data: conc
X-squared = 4.8444, df = 2, p-value = 0.0887
```

```
conc_matrix(c(50,47,56,5,14,8),ncol=3,byrow=T)
dimnames(conc)_list(c("recu","refuse"),
                    c("jury1","jury2","jury3"))
conc.chisq_chisq.test(conc)
print(conc.chisq)
```

Examen

Sortie SAS (test du Khi2)

TABLE OF A BY B

A	B			Total
Frequency	1	2	3	
Percent				
Row Pct				
Col Pct				
1	50	47	56	153
	27.78	26.11	31.11	85.00
	32.68	30.72	36.60	
	90.91	77.95	87.50	
2	5	14	8	27
	2.78	7.78	4.44	15.00
	18.52	51.85	29.63	
	9.09	22.95	12.50	
Total	55	61	64	180
	30.56	33.89	35.56	100.00

STATISTICS FOR TABLE OF A BY B

Statistic	DF	Value	Prob
Chi-Square	2	4.844	0.089
Likelihood Ratio Chi-Square	2	4.720	0.094
Mantel-Haenszel Chi-Square	1	0.179	0.672
Phi Coefficient		0.164	
Contingency Coefficient		0.162	
Cramer's V		0.164	

Sample Size = 180

Examen

Programme SAS (test du Khi2)

```
data concours;
do a=1 to 2;
  do b=1 to 3;
    input nb @@;
    output;
  end;
end;
cards;
50 47 56
5 14 8
;

proc freq data=concours order=data;
weight nb;
tables a*b /chisq ;
title 'test du khi2';
run;
```

PROBLÈME : Activité manuelle

Afin d'étudier l'influence d'un handicap sur la vitesse d'exécution d'une activité manuelle, un psychologue note le temps (en seconde) nécessaire à la réalisation d'un exercice manuel pour deux échantillons de personnes handicapées et non handicapées.

Non Handicapés	204	218	197	183	227	299	191	
Handicapés	243	228	261	202	343	242	220	239

Peut-on considérer que ces temps suivent la même répartition ?

Activité manuelle

Sortie SAS (test du Kolmogorov - Smirnov)

```
N P A R I W A Y P R O C E D U R E
Kolmogorov-Smirnov Test for Variable TEMPS
Classified by Variable TYPE
```

TYPE	N	EDF at maximum	Deviation from Mean at maximum
1	7	0.857142857	0.856719472
2	8	0.250000000	-.801387685
-----	-----	-----	-----
	15	0.533333333	

Maximum Deviation occurred at Observation 5
Value of TEMPS at maximum 227.000000

Kolmogorov-Smirnov 2-Sample Test (Asymptotic)
KS = 0.302896 D = 0.607143
KSa = 1.17311 Prob > KSa = 0.1275

N P A R I W A Y P R O C E D U R E

Cramer-von Mises Test for Variable TEMPS
Classified by Variable TYPE

TYPE	N	Summed Deviation from Mean
1	7	0.258835979
2	8	0.226481481

Cramer-von Mises Statistic (Asymptotic)
CM = 0.032354 CMa = 0.485317

Activité manuelle

Programme SAS (test du Kolmogorov - Smirnov)

```
data manuel;
input temps type;
cards;
204 1
218 1      proc print data=manuel;
197 1      run;
183 1      proc npar1way;
227 1      class type;
299 1      var temps;
191 1      run;
243 2
228 2
261 2
202 2
343 2
242 2
220 2
239 2
;
```

PROBLÈME : Exercice sur une variété de haricots

On croise une variété de haricots à graines blanches avec une variété à graines rouges (croisement $BbRr \times bbRR$). On sait que chez le haricot, la rencontre de 2 allèles dominants (B et R) donne des graines noires alors que la rencontre des 2 allèles récessifs (b et r) donne des graines blanches. En F₂ (2nde génération après le croisement), on observe 42 individus à graines noires, 19 individus à graines rouges et 15 individus à graines blanches. La ségrégation observée correspond-elle aux lois de Mendel ?