



BIOSTATISTIQUES AVEC R

CHEAT SHEETS DES FONCTIONS INDISPENSABLES



CLAIRE DELLA VEDOVA

A Lire – très important

*Le simple fait de lire le présent document vous donne le droit de *l'offrir en cadeau* à qui vous le souhaitez. Vous êtes autorisé à l'utiliser selon les mêmes conditions commercialement, c'est-à-dire à l'offrir sur votre blog, sur votre site web, à l'intégrer dans des packages et à l'offrir en bonus avec des produits, mais PAS à le vendre directement, ni à l'intégrer à des offres punies par la loi dans votre pays.*

Ce document est sous licence Creative Common 3.0 « Paternité – pas de modification », ce qui signifie que vous êtes libre de le distribuer à qui vous voulez, à condition de ne pas le modifier, et de toujours citer l'auteur Claire Della Vedova comme l'auteur de ce document, et d'inclure un lien vers : <https://statistique-et-logiciel-r.com>.



« Biostatistiques avec R – cheat sheet des fonctions indispensables » by Claire Della Vedova est mis à disposition selon les termes de la licence Creative Commons Paternité - Pas de Modification 3.0 Unported.

SOMMAIRE

1.	Paramètres Descriptifs	3
2.	Manipulation de data frame (tableau de données).....	5
3.	Evaluation de la normalité	8
4.	Evaluation de l'homogénéité	10
5.	Evaluation des outliers	12
6.	Comparaison de deux moyennes	14
6.1	Méthodes paramétriques	14
6.2	Méthodes non paramétriques	15
7.	Comparaison de K moyennes ($k > 2$).....	16
7.1	Méthodes paramétriques	16
7.2	Méthodes non paramétriques	18
8.	Fonctions utiles autour des pourcentages	20
9.	Comparaison d'un pourcentage a un pourcentage théorique	21
10.	Comparaison de deux pourcentages	22
10.1	Echantillons indépendants	22
10.2	Echantillons appariés	24
11.	Comparaison de k pourcentages ($k > 2$).....	25
12.	Comparaison de k pourcentages avec hypothèse d'ordre ($p_1 < \dots < p_k$)	27
13.	Corrélation.....	28

1. PARAMETRES DESCRIPTIFS

Paramètre	Exemple	Package	Commentaires	Infos complémentaires
Moyenne	<code>mean (iris\$Sepal.Length, na.rm=TRUE)</code>	base	na.rm=TRUE permet d'éviter un message d'erreur en cas de présence d'une donnée manquante (NA)	
Médiane	<code>median (iris\$Sepal.Length, na.rm=TRUE)</code>	base		
Mmin	<code>min (iris\$Sepal.Length, na.rm=TRUE)</code>	base		
Max	<code>max (iris\$Sepal.Length, na.rm=TRUE)</code>	base		
[min ; max]	<code>range (iris\$Sepal.Length, na.rm=TRUE)</code>	base		
1er quartile	<code>quantile (iris\$Sepal.Length, 0.25, na.rm=TRUE)</code>	base		
3ème quartile	<code>quantile (iris\$Sepal.Length, 0.75, na.rm=TRUE)</code>	base		
Ecart type	<code>sd (iris\$Sepal.Length, na.rm=TRUE)</code>	base		
Variance	<code>var (iris\$Sepal.Length, na.rm=TRUE)</code>	base		
Intervalle inter quartile	<code>IQR (iris\$Sepal.Length, na.rm=TRUE)</code>	base		
Moyenne tronquée	<code>mean (iris\$Sepal.Length, trim=0.1)</code>	base	(ici troncature de 10%)	
Moyenne géométrique	<code>geoMean (iris\$Sepal.Length,)</code>	EnvStats	ne supporte pas les NA	la moyenne géométrique peut aussi être obtenue avec cette commande : <code>exp(mean(log(x)))</code>
Ecart type géométrique	<code>geoSD (iris\$Sepal.Length))</code>		ne supporte pas les NA	l'écart type géométrique peut aussi être obtenue avec cette commande : <code>exp(sd(log(x)))</code>

Paramètre	Exemple	Package	Commentaires	Infos complémentaires
Multiples paramètres	<code>summary</code> (iris) <code>stat.desc</code> (iris)	base pastecs		
Asymétrie (skewness) et aplatissement (kurtosis)	<code>stat.desc</code> (iris, norm=TRUE)	pastecs	la fonction réalise aussi un test d'asymétrie et d'aplatissement l'aplatissement (plus d'infos dans la page d'aide de la fonction stats.desc)	
Multiples paramètres par groupe	<code>by</code> (iris, iris\$Species, summary) <code>stat.desc</code> (iris, norm=TRUE)	base pastecs	norm=TRUE permet d'avoir accès aux paramètres permettant d'évaluer la normalité	

2. MANIPULATION DE DATA FRAME (TABLEAU DE DONNEES)

Méthode	Exemple	Package	Infos complémentaires
Afficher les 10 premières lignes	<code>head (iris, 10)</code>	utils	
Afficher les 10 dernières lignes	<code>tail (iris, 10)</code>	utils	
Afficher 10 lignes aléatoires	<code>sample_n (iris, 10, replace=FALSE)</code>	dplyr	
Afficher la structure des données	<code>str (iris)</code> <code>summary (iris\$Sepal.Length)</code>	utils	
Afficher les noms des variables	<code>names (iris)</code>	base	
Sélectionner des colonnes	<code>library (dplyr)</code> <code>iris2 <- iris %>%</code> <code>select (Sepal.Length, Petal.Length)</code>	dplyr	
Sélectionner des lignes	<code>library (dplyr)</code> <code>setosa <- iris %>%</code> <code>filter (Species=="setosa")</code>	dplyr	

Méthode	Exemple	Package	Infos complémentaires
Sélectionner des lignes selon plusieurs critères	<pre>library(dplyr)</pre> <pre>set_vir <- iris %>% filter (Species %in% c("setosa","virginica"))</pre>	dplyr	<p>Le package dplyr peut parfois entrer en conflit avec d'autre package, et cela empêche certaines fonctions de fonctionner correctement.</p> <p>Pour contourner le problème il suffit de rajouter dplyr :: nom_fonction, ex : dplyr :: filter()</p>
Exclure une colonne	<pre>iris3 <- iris %>% select (- Petal.Length)</pre>		
Exclure une ligne	<pre>virgi_versi <- iris %>% filter (Species!="setosa")</pre>	dplyr	
Couper un data frame en liste selon les modalités d'une variable catégorielle	<pre>iris_ls <- split (iris, iris\$Species)</pre>	base	
Regroupe les éléments d'une liste dans un data frame	<pre>iris_df <- data.frame (Reduce (rbind, iris_ls))</pre>	base	
Recréer des levels à partir de levels existants	<pre>levels (iris\$Species)</pre> <pre>levels (iris\$Species)[c(1,2)] <- c("virgi_et_versi")</pre> <pre>levels (iris\$Species)</pre>	base	

Méthode	Exemple	Package	Infos complémentaires
Changer le niveau de référence	<pre>levels (iris\$Species)</pre> <pre>iris\$Species <- relevel (iris\$Species, ref= "versicolor")</pre> <pre>levels (iris\$Species)</pre>		
Ajouter une modalité	<pre>iris\$Species <- as.character (iris\$Species)</pre> <pre>iris\$Species [1:5] <- c("new_setosa")</pre> <pre>iris\$species <- as.factor (iris\$Species)</pre> <pre>head (iris,7)</pre>		
Calculer les sommes de colonnes	<pre>colSums (iris[, (1:4)])</pre>	base	
Calculer les moyennes de colonnes	<pre>colMeans (iris [, (1:4)])</pre>	base	
Calculer les sommes de lignes	<pre>rowSums (iris [, (1:4)])</pre>	base	
Calculer les moyennes de lignes	<pre>rowMeans (iris [, (1:4)])</pre>		

3. EVALUATION DE LA NORMALITE

Méthode	Exemple	Package	Commentaires	Infos complémentaires
Evaluation graphique de la normalité QQ plot ou droite d'Henry	<code>qqnorm (iris\$Sepal.Length)</code> <code>qqline (iris\$Sepal.Length)</code>	stats		
Histogramme	<code>ggplot (iris, aes (Sepal.Length)) + geom_histogram (bins=7)</code>	ggplot2	Alternatives : <code>hist(iris\$Sepal.Length, breaks=6)</code>	bin et breaks permettent d'augmenter ou de réduire le nombre de classes
Plot de densité	<code>ggplot (iris, aes (Sepal.Length)) + geom_density()</code>	ggplot2	Alternatives : <code>plot(density(iris\$Sepal.Length))</code>	
<u>Evaluation de la normalité par test</u>	<i>H0 : les données sont distribuées selon une loi Normale</i> <i>H1 : les données ne sont pas distribuées selon une loi Normale</i> <i>si pval < 0.05 alors la normalité est rejetée</i> <i>si pval > 0.05 alors la normalité est acceptée</i>			
Test de Shapiro-Wilk	<code>shapiro.test (iris\$Sepal.Length)</code> <code>shapiro.test (residuals (mod))</code>	stats		Evaluation de la normalité des résidus d'une régression
Test de Kolmogorov-Smirnov	<code>ks.test (iris\$Sepal.Length, "pnorm")</code>	stats		
Test de Cramer-Von Mises	<code>cvm.test (iris\$Sepal.Length)</code>	nortest		

Méthode	Exemple	Package	Commentaires	Infos complémentaires
Test de Lilliefors	<code>lillie.test (iris\$Sepal.Length)</code>	nortest		
Evaluation visuelle de la normalité de résidus	<pre>mod <- lm (Sepal.Length ~ Sepal.Width, data=iris) plot (mod, 2)</pre>	stats		

4. EVALUATION DE L'HOMOGENEITE

Méthode	Exemple	Package	Commentaires	Infos complémentaires
<u>Test de comparaison de variances</u>	<p>H_0 : les variances des k groupes sont égales (homogénéité)</p> <p>H_1 : les variances d'au moins deux groupes sont différentes (hétérogénéité)</p> <p>si $pval < 0.05$ alors l'homogénéité est rejetée</p> <p>si $pval > 0.05$ alors l'homogénéité est acceptée</p>			
<u>Comparaisons de 2 variances</u>				autre syntaxe : var.test (iris\$Sepal.Length, iris\$Sepal.Width)
Test F	var.test (Puromycin\$rate ~ Puromycin\$state)	stats		
<u>Comparaisons de k variances</u>				
Test de Levene	leveneTest (iris\$Sepal.Length, iris\$Species)	car		
Test de Bartlett	bartlett.test (iris\$Sepal.Length, iris\$Species)	stats	idem	
Test de Fligner	fligner.test (iris\$Sepal.Length, iris\$Species)	stats		

Méthode	Exemple	Package	Commentaires	Infos complémentaires
<u>Evaluation de l'homogénéité de résidus par test</u>	<p><i>H0 : les résidus du modèle de régression linéaire sont homogènes</i></p> <p><i>H1 : les résidus du modèle de régression linéaire sont hétérogènes</i></p> <p><i>si pval < 0.05 alors l'homogénéité est rejetée</i></p> <p><i>si pval > 0.05 alors l'homogénéité est acceptée</i></p>			
Test de Breush-Pagan	<pre>mod <- lm (Sepal.Length ~ Sepal.Width, data=iris) ncvTest (mod)</pre>		car	
	bptest (mod)	lmttest	idem	version studentisée du ncvTest (bpTest= plus robuste)
<u>Evaluation visuelle de l'homogénéité de résidus</u>	mod <- lm (Sepal.Length ~ Sepal.Width, data=iris)			

5. EVALUATION DES OUTLIERS

Méthode	Exemple	Package	Commentaires	Infos complémentaires
<u>Détection par boxplot</u>				
	<code>boxplot (mtcars\$wt)</code>	graphics	Est détectée comme outlier toute donnée située en dehors de l'intervalle $1.5 * \text{IQR}$.	IQR est l'intervalle inter quartile (la différence entre le premier et troisième quartile)
	<code>boxplot.stats (mtcars\$wt)\$out</code>	grDevices	permet d'obtenir les valeurs des données outliers	
	<code>which (mtcars\$wt %in% boxplot.stats (mtcars\$wt)\$out)</code>		permet de connaître l'indice (le numéro de ligne) des données outlier	
Par groupe	<code>boxplot (mtcars\$wt ~ as.factor (mtcars\$gear))</code>			
<u>Détection par test</u>				
	$H_0 : \text{la valeur la plus élevée n'est pas outlier}$			
	$H_1 : \text{la valeur la plus élevée est un outlier}$ si $pval < 0.05$ alors la donnée la plus élevée est outlier			
	si $pval > 0.05$ alors la donnée la plus élevée n'est pas outlier			
	<code>grubbs.test (mtcars\$wt)</code>	outliers	permet de tester la valeur la plus élevée	
	<code>grubbs.test (mtcars\$wt, opposite = TRUE)</code>		opposite=TRUE , permet de tester la valeur la plus faible	

Méthode	Exemple	Package	Commentaires	Infos complémentaires
<u>Détection sur des résidus</u>	<pre>mod <- lm (Sepal.Length ~ Sepal.Width, data=iris)</pre> <pre>outlierTest (mod,3)</pre>			

6. COMPARAISON DE DEUX MOYENNES

6.1 Méthodes paramétriques

Méthode	Exemple	Package	Infos complémentaires
<p><i>H₀ : les deux moyennes sont égales</i></p>			
	<p><i>H₁ : les deux moyennes sont différentes</i></p> <p><i>si pval < 0.05 alors les moyennes sont différentes</i></p> <p><i>si pval > 0.05 alors les moyennes sont égales</i></p>		
<p>Echantillons indépendants et variances égales</p>			
Test de Student	<code>t.test (Puromycin\$rate ~ Puromycin\$state, var.equal = TRUE)</code>	stats	<u>syntaxe alternative</u> : <code>t.test (data\$groupe1, data\$groupe2, var.equal=TRUE)</code>
<p>Echantillons indépendants et variances différentes</p>			
Test de Welch	<code>t.test (Puromycin\$rate ~ Puromycin\$state)</code>		
<p>Echantillons appariés</p>			
Test de Student apparié	<code>t.test (iris\$Sepal.Length,iris\$Petal.Length, paired=TRUE)</code>	stats	

6.2 Méthodes non paramétriques

Méthode	Exemple	Package	Infos complémentaires
<u>Echantillons indépendants</u>			
Test de Wilcoxon	<code>wilcox.test</code> (Puromycin\$rate ~ Puromycin\$state)	stats	alternative : <code>wilcox.test (data\$groupe1, data\$groupe2)</code>
<u>Echantillons appariés</u>			
Test de Wilcoxon	<code>wilcox.test</code> (Puromycin\$rate ~ Puromycin\$state, paired=TRUE)	stats	alternative : <code>wilcox.test (data\$groupe1, data\$groupe2)</code>

7. COMPARAISON DE K MOYENNES (K > 2)

7.1 Méthodes paramétriques

Méthode	Exemple	Package	Infos complémentaires
<p><i>H0 : les k moyennes sont égales</i></p>			
<p><i>H1 : au moins 2 moyennes sont différentes</i></p>			
<p><i>si pval < 0.05 alors au moins 2 moyennes sont différentes</i></p>			
<p><i>si pval > 0.05 alors les k moyennes sont égales</i></p>			
ANOVA	<pre>my_lm <- lm (Sepal.Length ~ Species, data=iris) anova (my_lm)</pre>	stats	
<pre>my_aov <- aov (Sepal.Length ~ Species, data=iris) summary (my_aov)</pre>			
<p><u>Comparaisons multiples après ANOVA</u></p>			
<p>Comparaisons 2 à 2</p>			
Test de Tukey	<pre>pairwise.t.test (iris\$Sepal.Length, iris\$Species, "holm")</pre>	stats	ajustement des pval selon "Holm"
<p>Options dispo: "bonferroni", "holm", "hochberg", "hommel", "BH", "fdr", "BY" "none" (cf. ?p.adjust)</p>			
ou	<pre>lsmeans (my_aov, "Species")</pre>	lsmeans	Ne fonctionne pas sur un objet lm

<i>Méthode</i>	<i>Exemple</i>	<i>Package Infos complémentaires</i>
	ou	
	<pre>TukeyHSD (my_aov) plot (TukeyHSD(my_aov)) mc <- glht (my_lm, linfct = mcp(Species = "Tukey")) summary(mc) plot (mc)</pre>	multcomp
Comparaisons au Témoin		
Test de Dunnett	<pre>mc <- glht (my_lm, linfct = mcp(Species = "Dunnet")) summary(mc) plot(mc)</pre>	multcomp

7.2 Méthodes non paramétriques

Méthode	Exemple	Package	Infos complémentaires
	<p>H_0 : les k moyennes sont égales</p> <p>H_1 : au moins 2 moyennes sont différentes</p> <p>si $pval < 0.05$ alors au moins 2 moyennes sont différentes</p> <p>si $pval > 0.05$ alors les k moyennes sont égales</p>		
Test de Kruskal-Wallis	<code>kruskal.test</code> (Sepal.Length ~ Species, data=iris)	stats	Test de Kruskal-Wallis
<u>Comparaisons multiples après Kruskall-Wallis</u>			
Comparaisons 2 à 2	<code>pairwise.t.test</code> (iris\$Sepal.Length, iris\$Species, "holm")	stats	ajustement des pval selon "Holm" Options dispo: "bonferroni", "holm", "hochberg", "hommel", "BH", "fdr", "BY" "none" (cf. ?p.adjust)
Toutes les comparaisons des moyennes à la moyenne du témoin	recopier les pvalues d'intérêt à partir de <code>pairwise.wilcox.test</code> sans correction (method="none") <code>pairwise.wilcox.test</code> (iris\$Sepal.Length, iris\$Species,"none") <code>my_pval <- c(5.9e-07 ,8.3e-14)</code> <code>my_pval_ajust <- p.adjust(my_pval, "holm")</code>		

<i>Méthode</i>	<i>Exemple</i>	<i>Package Infos complémentaires</i>
<u>FONCTIONS UTILES</u>		
Changer la modalité référence pour l'ANOVA ou le test de Dunnett	iris\$Species <- relevel (iris\$Species, ref = "virginica")	
Calculer l'intervalle de confiance d'une moyenne par bootstrap		
Globalement	slipper_ci (iris, mean(Sepal.Length), B=1000, lower=0.025, upper=0.975)	tidyverse, slipper
Par groupe	iris %>% filter (Species=="setosa") %>% slipper_ci (mean(Sepal.Length), B=100, lower=0.025, upper=0.975)	

8. FONCTIONS UTILES AUTOUR DES POURCENTAGES

Méthode	Exemple	Package	Infos complémentaires
TABLE DE CONTINGENCE	<code>table</code> (Melanoma\$sex, Melanoma\$ulcer) ou <code>xtabs</code> (~ Melanoma\$sex + Melanoma\$ulcer)	base stats	Melanoma est un dataset du package MASS
CALCUL DE POURCENTAGES			
pourcentages globaux	<code>TC <- table</code> (Melanoma\$sex, Melanoma\$ulcer) <code>prop.table</code> (TC)	base	
pourcentages par ligne	<code>prop.table</code> (TC,1)	base	
pourcentages par colonne	<code>prop.table</code> (TC,2)	base	
CALCUL DE L'INTERVALLE DE CONFIANCE D'UN POURCENTAGE	<code>prop.test</code> (90, 90+115)\$conf.int	stats	<code>prop.test</code> (succes, nb_Total)\$conf.int

9. COMPARAISON D'UN POURCENTAGE A UN POURCENTAGE THEORIQUE

Méthode	Exemple	Package	Infos complémentaires
	<p>H_0 : le pourcentage est égal au pourcentage théorique</p> <p>H_1 : les deux pourcentages sont différents</p> <p>si $pval < 0.05$ alors les 2 pourcentages sont différents</p> <p>si $pval > 0.05$ alors les 2 pourcentages sont égaux</p>		
ici p théorique=0.5	<code>binom.test (55,100,0.5)</code> ou <code>binom.test (c (55,45),0.5)</code>	stats	<code>binom.test (Nb_succès, Nb Total, proba_théorique)</code> <u>ou :</u> <code>binom.test c (Nb_succès , Nb_echecs), proba_theorique)</code>

10. COMPARAISON DE DEUX POURCENTAGES

10.1 Echantillons indépendants

Méthode	Exemple	Package	Infos complémentaires
	<p>H_0 : les deux pourcentages sont égaux</p> <p>H_1 : les deux pourcentages sont égaux</p> <p>si $pval < 0.05$ alors les pourcentages sont différents</p> <p>si $pval > 0.05$ alors les pourcentages sont égaux</p>		
Test du Chi2	<pre>TC <- table (Melanoma\$sex, Melanoma\$ulcer) chisq.test (TC, correct = FALSE)</pre>	stats	Affiche le warning suivant si au moins un effectif calculé est autre synthaxe : <code>prop.test(TC)</code> la correction de continuité est l'option par défaut
Test du Chi2 avec correction de continuité	<code>chisq.test (TC)</code>		à utiliser si au moins un effectif calculé est <5
Test de Fisher	<code>fisher.test (TC)</code>		

<i>Méthode</i>	<i>Exemple</i>	<i>Package</i>	<i>Infos complémentaires</i>
<u>COMPARAISONS MULTIPLES</u>			
Comparaisons multiples (toutes les comparaisons 2 à 2)	<code>pairwise.prop.test (TC, "holm")</code> <code>pairwise.fisher.test (TC, "holm")</code>	stats	Options dispo: "bonferroni", "holm", "hochberg", "hommel", "BH", "fdr", "BY" "none"

10.2 Echantillons appariés

Méthode	Exemple	Package	Infos complémentaires
	<p><i>H0 : les deux pourcentages sont égaux</i></p> <p><i>H1 : les deux pourcentages sont égaux</i></p> <p><i>si pval < 0.05 alors les pourcentages sont différents</i></p> <p><i>si pval > 0.05 alors les pourcentages sont égaux</i></p>		
Test du Chi2 de Mac Nemar avec correction de continuité	<pre>TC <- matrix(c(794, 86, 150, 570), nrow = 2, dimnames = list("Test1" = c("ok", "pas_ok"), "Test2" = c("ok", "pas ok"))) mcnemar.test(TC)</pre>	stats	la correction de continuité est l'option par défaut. A utiliser si au moins un effectif calculé est <5
Test du Chi2 de Mac	<code>mcnemar.test(TC, correct=FALSE)</code>	stats	

11. COMPARAISON DE K POURCENTAGES (K >2)

Méthode	Exemple	Package	Infos complémentaires
	<p>H_0 : les k pourcentages sont égaux</p> <p>H_1 : au moins deux pourcentages sont différents</p> <p>si $pval < 0.05$ alors les pourcentages sont différents</p> <p>si $pval > 0.05$ alors les pourcentages sont égaux</p>		
Test du Chi2	<pre>TC <- table(Melanoma\$sex, Melanoma\$ulcer)</pre> <pre>chisq.test(TC, correct = FALSE)</pre>	stats	<p>Affiche le warning suivant si au moins un effectif calculé est <5 : "In chisq.test (my.mat) : l'approximation du Chi-2 est peut-être incorrecte"</p> <p>autre syntaxe : prop.test (TC)</p>
Test du Chi2 avec correction de continuité	chisq.test (TC)		à utiliser si au moins un effectif calculé est <5
Test de Fisher	fisher.test (TC)		à utiliser si au moins un effectif calculé est <5
<u>COMPARAISONS MULTIPLES</u>			
Comparaisons multiples (toutes les comparaisons 2 à 2)	<pre>pairwise.prop.test(TC, "holm")</pre> ou <pre>pairwise.fisher.test(TC, "holm")</pre>		Options dispo: "bonferroni", "holm", "hochberg", "hommel", "BH", "fdr", "BY", "none"

<i>Méthode</i>	<i>Exemple</i>	<i>Package</i>	<i>Infos complémentaires</i>
Comparaisons multiples (toutes les comparaisons au témoin)	<pre>pairwise.prop.test (TC, "none")</pre> <pre>my_pval <- c(1.6e-06 , 0.42) my_pval_ajust <- p.adjust (my_pval, "holm")</pre> <pre>my_pval_ajust</pre>	stats	

12. COMPARAISON DE K POURCENTAGES AVEC HYPOTHESE D'ORDRE (P1 < ...<PK)

Méthode	Exemple	Package	Infos complémentaires
	<p>H_0 : les pourcentages sont égaux ($p_1 = p_2 = \dots = p_n$)</p> <p>H_1 : les pourcentages sont ordonnés ($p_1 \leq p_2 \leq \dots \leq p_j$ ou $p_1 \geq p_2 \geq \dots \geq p_j$ avec au moins une inégalité stricte)</p> <p>si $pval < 0.05$ alors les pourcentages sont globalement ordonnés</p> <p>si $pval > 0.05$ alors les pourcentages sont globalement égaux</p>		<pre>succes <- c(90,65,68,10) total <- rep(100,4) prop.trend.test (succes,total)</pre>

13. CORRELATION

Méthode	Exemple	Package	Infos complémentaires
Evaluation de la relation monotone (Spearman & Kendall) ou linéaire (Pearson) entre deux variables continues	<p>H0 : pas de dépendance monotone/linéaire entre les deux variables</p> <p>H1 : dépendance monotone/linéaire entre les deux variables</p> <p>si pval <0.05 alors il existe une monotone/linéaire entre les deux variables</p> <p>si pval > 0.05 alors il n'existe pas de dépendance monotone/linéaire entre les deux variables</p>		
Coefficient de corrélation de Pearson	<code>cor.test (iris\$Sepal.Length, iris\$Petal.Length , method = "pearson")</code>	stats	la méthode par défaut ="Pearson"
Coefficient de corrélation de Spearman	<code>cor.test (iris\$Sepal.Length, iris\$Petal.Length, method = "spearman")</code>	stats	
Coefficient de corrélation de Kendall	<code>cor.test (iris\$Sepal.Length, iris\$Petal.Length, method = "kendall")</code>	stats	
Matrices			
Matrice de corrélation selon Pearson	<code>cor (mtcars, method = "pearson")</code>	stats	

Méthode	Exemple	Package	Infos complémentaires
Matrice de corrélation selon Spearman	<code>cor (mtcars, method = "spearman")</code>	stats	
Matrice de corrélation selon Kendall	<code>cor (mtcars, method="kendall")</code>	stats	
Matrice de variance covariance	<code>cov (mtcars)</code>		
<u>Test d'autocorrélation sur les résidus</u>	H0 : les résidus ne sont pas auto-corrélées H1 : les résidus sont auto-corrélées si pval <0.05 alors il existe une auto-corrélation si pval > 0.05 alors il n'existe pas d'auto-corrélation		
Test de Durbin-Watson	<code>mod <- lm (Sepal.Length~Petal.Length, data=iris)</code> <code>durbinWatsonTest (mod)</code>	car	

Crédits photos : Lukas