

Ensemble de fonctions pour les tests statistiques courants en R

PAPA B. NDIAYE

2024-08-22

```
#####  
# Charger les bibliothèques nécessaires####  
#####  
library(pacman)  
p_load(dplyr)  
p_load(stats)  
p_load(car)  
p_load(MASS)  
p_load(coin)  
  
#####  
# Définition des fonctions#####  
#####  
  
test_student <- function(group1, group2) {  
  cat("\nTest t de Student\n")  
  tryCatch({  
    p1 <- shapiro.test(group1)$p.value  
    p2 <- shapiro.test(group2)$p.value  
    if (p1 < 0.05 || p2 < 0.05) {  
      cat("Attention : Les données ne semblent pas  
        suivre une distribution normale.\n")  
      cat("Considérez l'utilisation d'un test  
        non paramétrique comme le test de Mann-Whitney U.\n")  
    }  
    p_var <- var.test(group1, group2)$p.value  
    if (p_var < 0.05) {  
      cat("Attention :  
        Les variances des deux groupes semblent différentes.\n")  
      cat("Considérez l'utilisation du test de Welch.\n")  
    }  
    t_result <- t.test(group1, group2, var.equal = (p_var >= 0.05))  
    cat("t = ", t_result$statistic, ", p = ", t_result$p.value, "\n")  
  }, error = function(e) {  
    cat("Erreur dans le test t de Student : ", e$message, "\n")  
  })  
}  
  
# Fonction pour le test de Mann-Whitney  
test_mannwhitney <- function(group1, group2) {  
  cat("\nTest de Mann-Whitney U\n")  
}
```

```

tryCatch({
  result <- wilcox.test(group1, group2, exact = FALSE)
  cat("stat = ", result$statistic, ", p = ", result$p.value, "\n")
}, error = function(e) {
  cat("Erreur dans le test de Mann-Whitney : ", e$message, "\n")
})
}

test_anova <- function(data, formula) {
  cat("\nANOVA\n")
  tryCatch({
    model <- lm(as.formula(formula), data = data)
    anova_result <- anova(model)
    print(anova_result)
    p_resid <- shapiro.test(residuals(model))$p.value
    if (p_resid < 0.05) {
      cat("Attention :
          Les résidus ne semblent pas suivre une distribution normale.\n")
    }
    independent_var <- all.vars(as.formula(formula))[2]
    dependent_var <- all.vars(as.formula(formula))[1]
    p_var <- leveneTest(as.formula(paste(dependent_var, "~", independent_var)),
                        data = data)$'Pr(>F)'[1]
    if (p_var < 0.05) {
      cat("Attention :
          Les variances des groupes semblent différentes.\n")
    }
  }, error = function(e) {
    cat("Erreur dans l'ANOVA : ", e$message, "\n")
  })
}

# Fonction pour le test de Kruskal-Wallis
test_kruskal <- function(...) {
  cat("\nTest de Kruskal-Wallis\n")
  tryCatch({
    result <- kruskal.test(list(...))
    cat("H = ", result$statistic, ", p = ", result$p.value, "\n")
  }, error = function(e) {
    cat("Erreur dans le test de Kruskal-Wallis : ", e$message, "\n")
  })
}

# Fonction pour le test du chi-carré
test_chi2 <- function(data, var1, var2) {
  cat("\nTest du chi-carré\n")
  tryCatch({
    contingency_table <- table(data[[var1]], data[[var2]])
    chi2_result <- chisq.test(contingency_table)
    if (any(chi2_result$expected < 5)) {
      cat("Attention : Certains effectifs théoriques sont
          inférieurs à 5, les résultats peuvent ne pas être fiables.\n")
    }
  }, error = function(e) {
    cat("Erreur dans le test du chi-carré : ", e$message, "\n")
  })
}

```

```

    }
    cat("chi2 = ", chi2_result$statistic,
        ", p = ", chi2_result$p.value, "\n")
}, error = function(e) {
    cat("Erreur dans le test du chi-carré : ", e$message, "\n")
})
}

# Fonction pour le test exact de Fisher
test_fisher <- function(data, var1, var2) {
    cat("\nTest exact de Fisher\n")
    tryCatch({
        contingency_table <- table(data[[var1]], data[[var2]])
        fisher_result <- fisher.test(contingency_table)
        cat("odds ratio = ", fisher_result$estimate,
            ", p = ", fisher_result$p.value, "\n")
    }, error = function(e) {
        cat("Erreur dans le test exact de Fisher : ", e$message, "\n")
    })
}

# Fonction pour la corrélation de Pearson
test_pearson <- function(x, y) {
    cat("\nCorrélation de Pearson\n")
    tryCatch({
        p1 <- shapiro.test(x)$p.value
        p2 <- shapiro.test(y)$p.value
        if (p1 < 0.05 || p2 < 0.05) {
            cat("Attention :
                Les données ne semblent pas suivre une distribution normale.\n")
            cat("Considérez l'utilisation de la corrélation de Spearman.\n")
        }
        result <- cor.test(x, y, method = "pearson")
        cat("r = ", result$estimate, ", p = ", result$p.value, "\n")
    }, error = function(e) {
        cat("Erreur dans la corrélation de Pearson : ", e$message, "\n")
    })
}

# Fonction pour la corrélation de Spearman
test_spearman <- function(x, y) {
    cat("\nCorrélation de Spearman\n")
    tryCatch({
        result <- cor.test(x, y, method = "spearman")
        cat("rho = ", result$estimate, ", p = ", result$p.value, "\n")
    }, error = function(e) {
        cat("Erreur dans la corrélation de Spearman : ", e$message, "\n")
    })
}

# Tester les fonctions avec des exemples de données
set.seed(123)

```

```
# Exemples pour les tests t de Student et Mann-Whitney
group1 <- rnorm(50, mean = 5, sd = 2)
group2 <- rnorm(50, mean = 7, sd = 2)
test_student(group1, group2)
```

```
##
## Test t de Student
## t = -6.07178 , p = 2.403003e-08
```

```
test_mannwhitney(group1, group2)
```

```
##
## Test de Mann-Whitney U
## stat = 491 , p = 1.70464e-07
```

```
# Exemples pour les tests ANOVA et Kruskal-Wallis
data <- data.frame(
  y = c(group1, group2, rnorm(50, mean = 6, sd = 2)),
  x = factor(rep(1:3, each = 50))
)
test_anova(data, "y ~ x")
```

```
##
## ANOVA
## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value    Pr(>F)
## x           2  139.46   69.731   19.692 2.648e-08 ***
## Residuals 147  520.55    3.541
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
test_kruskal(group1, group2, rnorm(50, mean = 6, sd = 2))
```

```
##
## Test de Kruskal-Wallis
## H = 30.35692 , p = 2.559053e-07
```

```
# Exemples pour les tests du chi-carré et exact de Fisher
data_chi <- data.frame(
  var1 = sample(c("A", "B"), 100, replace = TRUE),
  var2 = sample(c("X", "Y"), 100, replace = TRUE)
)
test_chi2(data_chi, "var1", "var2")
```

```
##
## Test du chi-carré
## chi2 = 0.6527948 , p = 0.4191152
```

```
test_fisher(data_chi, "var1", "var2")
```

```
##  
## Test exact de Fisher  
## odds ratio = 1.499842 , p = 0.4192918
```

```
# Exemples pour les corrélations de Pearson et Spearman  
x <- rnorm(100)  
y <- x + rnorm(100)  
test_pearson(x, y)
```

```
##  
## Corrélation de Pearson  
## r = 0.717782 , p = 4.338256e-17
```

```
test_spearman(x, y)
```

```
##  
## Corrélation de Spearman  
## rho = 0.6882928 , p = 0
```