# Tests d'ajustement fondés sur la méthode Monte Carlo randomisée pour des distributions exponentielles

Jean-Marie Dufour* et Abdeljelil Farhat**

*McGill University, Leacock Building, Room 519, 855 Sherbrooke Street West,
Montréal, Québec H3A 2T7, Canada.
jean-marie.dufour@mcgill.ca,
http ://www.jeanmariedufour.com
**Unité de recherche : Économie Appliquée et Simulation Mahdia,
Faculté des sciences économiques et de gestion de Mahdia,
Université de Monastir, Tunisie;
abdeljelil.farhat@fsegma.rnu.tn

**Résumé.** La distribution exponentielle est largement utilisée pour la modélisation de données sur la durée des événements statistiques, en statistique, en économétrie et en finance. Ainsi, les tests d'exponentialité sont un important problème lorsque l'on s'intéresse à l'étude des données. Cependant, la plupart des tests proposés sont limités à la distribution exponentielle avec un seul paramètre et les tableaux de valeurs critiques sont disponibles uniquement pour un nombre limité de tailles d'échantillons. Ceci peut être la source d'importants problèmes de niveaux et de puissances des tests.

Dans cette étude, nous proposons d'abord l'utilisation de la technique des tests de Monte Carlo randomisés, en vue de contrôler la taille de différents tests d'exponentialité comportant un paramìre d'échelle et un paramètre de localisation. Nous montrons que les tests obtenus de cette façon sont exacts pour toute taille d'échantillon. Nous proposons aussi des modifications de la procédure fondées sur des estimateurs de moments et nous montrons que ces modifications améliorent sensiblement les puissances de plusieurs tests proposés dans la littérature antérieure.

Cette étude est achevée par la proposition de nouveaux tests basés sur les procédures des combinaisons de plusieurs statistiques de tests. Ces tests ont montré de très bonnes propriétés de puissance.

## 1 Introduction

Dans cette étude, nous proposons d'abord l'utilisation de la technique des tests de Monte Carlo Randomisée (MCR) (voir Dufour (2006)), en vue de contrôler le niveau des différents tests d'exponentialité. Nous montrons aussi sur le plan théorique et par simulation que les tests obtenus de cette façon sont exacts pour toute taille d'échantillon. Nous proposons aussi des modifications de la procédure fondée sur les estimateurs des moments et nous montrons que ces modifications améliorent sensiblement les puissances de plusieurs tests proposés dans la littérature antérieure.

En résumé, les principaux objectifs de ce travail sont les suivants :

(i) étendre l'utilisation des tests de MCR pour effectuer des tests d'exponentialité ;

(ii) montrer que les tests de MCR ont toujours un niveau exact, tandis que les tests originaux ont souvent un niveau biaisé ;

(iii) proposer des modifications des estimateurs utilisés et des procédures de combinaison fondées sur les propriétés des tests de MCR pour obtenir des tests plus performants dans le cas où la taille d'échantillon est réduite ;

(iv) effectuer la comparaison des performances du point de vue de la puissance pour différents tests afin de sélectionner les meilleurs selon plusieurs contre-hypothèses courantes.

Après une présentation de la problématique dans la seconde section, nous discutons la procédure des tests de MCR dans la section 3. La section 4 est réservée à la présentation des statistiques de test considérées ainsi qu'à la procédure des combinaisons proposées. La section 5 est consacrée à la discussion des résultats obtenus dans nos expériences de simulation. Dans une dernière section, nous énonçons quelques conclusions.

## 2 Problématique

Pour confronter ce sujet, considérons un échantillon aléatoire $X_1$, ..., $X_n$ associé à une variable aléatoire $X$ dont la fonction de répartition (f.r.) est $F(x;\alpha,\beta)$, où $\alpha \in R$ et $\beta \in R^+$ sont, respectivement, deux paramètres inconnus de position et d'échelle. Le problème, dans cette étude, est de tester

$$H_0 : F(x;\alpha,\beta) = F_0\left(\frac{x-\alpha}{\beta}\right) = 1 - exp\left\{-\frac{(x-\alpha)}{\beta}\right\},\ x > \alpha$$

contre l'alternative

$$H_1 : F(x;\alpha,\beta) \neq F_0\left(\frac{x-\alpha}{\beta}\right),\ \text{ pour au moins une valeur de } x,$$

où $F_0$ est la f.r. de la distribution exponentielle ($E$).

Si $\alpha$ et $\beta$ ont été connus, il serait équivalent à vérifier si $Y_i = (X_i - \alpha)/\beta,\ 1 \leq i \leq n$, sont régis par la f.r. $F_0(x)$. Mais, étant donné qu'ils ne sont pas connus, $\alpha$ et $\beta$ seront remplacés par des estimateurs $\hat{\alpha}_X = \hat{\alpha}(X_1, ..., X_n)$ et $\hat{\beta}_X = \hat{\beta}(X_1, ..., X_n)$. Les statistiques que nous examinerons seront des fonctions des variables observables $W_i = (X_i - \hat{\alpha}_X)/\hat{\beta}_X,\ 1 \leq i \leq n$. Ces variables ne sont pas mutuellement indépendantes, mais leur distribution est libre des paramètres inconnus $\alpha$ et $\beta$. Cela découle du fait que, si on note $\hat{\alpha}_Y = \hat{\alpha}(Y_1, ..., Y_n)$ et $\hat{\beta}_Y = \hat{\beta}(Y_1, ..., Y_n)$, nous avons

$$\frac{Y_i - \hat{\alpha}_Y}{\hat{\beta}_Y} = \frac{(X_i - \alpha)/\beta - (\hat{\alpha}_X - \alpha)/\beta}{\hat{\beta}_X/\beta} = \frac{X_i - \hat{\alpha}_X}{\hat{\beta}_X},\ 1 \leq i \leq n.$$

Par conséquent, toute statistique construite à partir des variables $W_i$ est pivotale. Nous rappelons qu'une statistique est dite pivotale pour un paramètre $\theta$ si sa loi est connue et indépendante de $\theta$.

Nous savons aussi que, pour la majorité des tests d'ajustement (goodness-of-fit tests (GOFT)), la distribution de la statistique de test, sous l'hypothèse nulle, est très compliquée ou même parfois inconnue. Il s'ensuit que les méthodes habituelles sont limitées, soit à des procédures asymptotiques dont la performance n'est pas très fiable, ou d'exiger la construction de tables spécialisées (voir, par exemple, D'Agostino et Stephens (1986)). Pour la majorité, ces tables ont été obtenues par simulation et leur utilisation génère des problèmes selon la nature des données observées. En outre, les tableaux ne peuvent pas couvrir toutes les tailles d'échantillon (voir, par exemple, Lilliefors (1969) ou Maurer et Margolin (1976)).

Pour résoudre ce problème, nous optons pour la mise en place des procédures basées sur les méthodes MCR initialement proposées par Dwass (1957) et Barnard (1963) et approfondies par Dufour (2006). Nous signalons que ces procédures ont été utilisées par Birnbaum (1974), Dufour et al. (1998), Dufour et Farhat (2002) et Dufour et al. (2004)].

Pour l'estimation des paramètres inconnus $\alpha$ et $\beta$, Stephens (1974) a souvent fait recours aux estimateurs du maximum de vraisemblance (Maximum Likelihood Estimates (MLE)). En dépit des qualités reconnues de ces estimateurs, nous notons en cette recherche, où les tailles d'échantillons sont réduites, que l'utilisation des estimateurs des moments améliore clairement, dans plusieurs situations, la puissance des tests étudiés. Cela nous a permis aussi d'introduire des tests de MCR combinés qui ont prouvé, pour la majorité des cas, d'être nettement plus puissants que les tests initiaux.

# 3 La procédure des tests de Monte Carlo randomisés

Soit $S$ une statistique de test pivotale dont la distribution est continue sous l'hypothèse nulle $H_0$. On rejette l'hypothèse $H_0$ lorsque $S$ est "grand" : $S \geq c$, où $c$ dépend du niveau $\alpha$ du test. On peut réécrire ce test sous la forme : $G(S) \leq \alpha$, où

$$G(x) = P[S \ \geq \ x \mid H_0].$$

En général, la fonction $G(x)$ est inconnue. Nous allons toutefois supposer que, sous l'hypothèse nulle, on peut engendrer $N$ réalisations indépendantes ou, à la rigueur interchangeables, $S_1, S_2, ..., S_N$ de la statistique $S$ [voir Dwass (1957) et Dufour (2006)]. Rappelons, d'abord, la définition du concept d'interchangeabilité :

**Définition :** On dit qu'un vecteur aléatoire $S_{(N)} = (S_1, ..., S_N)'$ est à composantes interchangeables si et seulement si sa loi conjointe demeure invariante par rapport à toutes les permutations de ses composantes.

D'après cette définition, on peut déduire que des variables aléatoires interchangeables sont forcément équidistribuées. Pour l'application de la technique des tests de MCR, on définit

$$\hat{G}_N(x) = \frac{1}{N}\sum_{i=1}^{N} I_{[0,\infty)}(S_i - x), \text{ où } I_A(x) = \begin{cases} 1, & \text{si } x \in A, \\ 0, & \text{si } x \notin A, \end{cases}$$

et la fonction critique empirique

$$\hat{p}_N(x) = \frac{N\hat{G}_N(x) + 1}{N + 1}. \tag{1}$$

Ainsi, la région critique de niveau $\alpha$ associée à un test de MCR s'exprime

$$\hat{p}_N(S_0) \leq \alpha.$$

On peut montrer que, peu importe $0 \leq \alpha \leq 1$,

$$P[\hat{p}_N(S_0) \leq \alpha | H_0] = \frac{I[\alpha(N+1)]}{N+1} \tag{2}$$

où $I[x]$ est la fonction partie entière et où $N$ est un entier choisi de façon à ce que $\alpha N$ soit entier, pour un test unilatéral de niveau $\alpha$ [voir Dufour (2006)].

Dans le cas d'une statistique dont la distribution n'est pas continue (sous l'hypothèse nulle), il se peut qu'on rencontre des ex aequo parmi $S_0, S_1, ..., S_N$. Dans une telle situation le test n'est plus exact, dans le sens que l'égalité (2) n'est plus toujours vraie. Pour corriger ce problème, on peut utiliser la technique de randomisation suivante [voir Dufour (2006), section 2.2] : considérer ($N$+1) réalisations indépendantes $U_0, U_1, ..., U_N$ d'une uniforme U(0,1), pour j = 0, 1, ..., $N$, associer $S_j$ à $U_j$ et enfin affirmer que

$$(S_i, U_i) > (S_j, U_j) \Longleftrightarrow S_i > S_j \text{ ou } \{S_i = S_j \text{ et } U_i > U_j\}.$$

Comme dans le cas d'une statistique de test de loi continue, on peut définir ainsi une région critique randomisée :

$$\tilde{p}_N(S_0) \leq \alpha$$

où

$$\tilde{p}_N(x) = \frac{N\tilde{G}_N(x) + 1}{N+1} \tag{3}$$

et

$$\tilde{G}_N(x) = 1 - \frac{1}{N}\sum_{i=1}^{N} I_{[0,\infty)}(x - S_i) + \frac{1}{N}\sum_{i=1}^{N} I_{[0]}(S_i - x) I_{[0,\infty)}(U_i - U_0).$$

Dans Dufour (2006, section 2), on montre que ces tests basés sur les méthodes randomisées sont des test exacts dans le sens que la région critique randomisée a le même niveau que la région critique $G(S_0) \leq \alpha$. Plus précisément, peu importe $0 \leq \alpha \leq 1$ :

$$P[\hat{p}_N(S_0) \leq \alpha | H_0] \leq P[\tilde{p}_N(S_0) \leq \alpha | H_0] = \frac{I[\alpha(N+1)]}{N+1}.$$

Les procédures des tests de MCR présentées dans cette section pourront être utilisées pour effectuer des tests d'ajustement, d'indépendance, pour détecter des données aberrantes, etc. Il faut et il suffit de connaître sous l'hypothèse nulle la distribution des observations ou, dans le cas de modèles de régression, la distribution des résidus. Si la statistique de test est pivotale, on peut se permettre de ne connaître sous $H_0$ la loi des observations qu'à un paramètre de position et à un paramètre d'échelle près.

# 4 Théorie

Pour effectuer des tests de l'hypothèse nulle $H_0$, nous considérons les statistiques de test suivantes.

## 4.1 Tests basées sur la fonction de répartition empirique

Les tests basées sur la fonction de répartition empirique (f.r.e.) mesurent des distances entre la f.r.e. et la f.r. de la distribution sous l'hypothèse nulle $H_0$. Les plus connues sont celui de Kolmogorov-Smirnov ($KS$) [Kolmogorov (1933), Smirnov (1939)], le test de Cramér-von Mises ($CM$) [Cramér (1928)], le test d'Anderson-Darling ($AD$) [Anderson et Darling (1954)] et le test de Watson ($W$) [Watson (1961)]. Pour un échantillon fini, les distributions des statistiques de $AD$ et $CM$ sont assez compliquées, mais une théorie asymptotique est disponible. Pour la statistique de $KS$, les distribution exacte et limite et même la distribution asymptotique sont non standard. Donc, les points critiques doivent être estimés.

Soient $X_{(1)} < ... < X_{(n)}$ $n$ observations exprimées dans un ordre croissant et soient $W_i = (X_i - \hat{\alpha}_X)/\hat{\beta}_X,\ 1 \leq i \leq n$ leurs transformations avec $W_{(1)} < ... < W_{(n)}$ leur énumération dans l'ordre croissant. En outre, notons

$$\begin{aligned}
F_0(x) &= \begin{cases} 0, & \text{si} \quad x < W_{(1)} \\ i/n, & \text{si} \quad W_{(i)} \leq x < W_{(i+1)}, \quad i = 1, ..., n-1, \\ 1, & \text{si} \quad W_{(n)} \leq x \end{cases} \\
Z_i &= F_0(W_{(i)}), i = 1, ..., n, \\
D^+ &= max_{i=1,...,n}\left(\frac{i}{n} - Z_i\right) \text{ et} \\
D^- &= max_{i=1,...,n}\left(Z_i - \frac{i-1}{n}\right).
\end{aligned}$$

Les statistiques sont définies comme suit :

$$\begin{aligned}
KS &= max(D^+, D^-), \\
CM &= \frac{1}{12n} + \sum_{i=1}^{n}\left[Z_i - \frac{2i-1}{2n}\right]^2, \\
AD &= -n - \frac{1}{n}\sum_{i=1}^{n}(2i-1)\left[Log(Z_i) + Log(1 - Z_{n+1-i})\right], \\
W &= n\int_{-\infty}^{\infty}\{F_n(x) - F_0(x) - \int_{-\infty}^{\infty}[F_n(x) - F_0(x)]dF_0(x)\}^2 dF_0(x) \\
&= CM - n(\bar{Z} - 0,5)^2.
\end{aligned}$$

Pour effectuer le test de l'hypothèse nulle $H_0$, nous procédons comme suit [ (D'Agostino et Stephens, 1986, page 141)] :

a) calculer les estimateurs $\hat{\beta}_X = n(\bar{X} - X_{(1)})/(n-1)$ et $\hat{\alpha}_X = X_{(1)} - \hat{\beta}_X/n$,

b) calculer $W_{(i)} = (X_{(i)}$ -$\hat{\alpha}_X)/\hat{\beta}_X, 1 \leq i \leq n,$

c) calculer $Z_i = F_0(W_{(i)}) = 1 - exp(-W_{(i)}), 1 \leq i \leq n,$

d) calculer les statistiques $KS, CM, AD$ et $W$.

Pour ces quatre premières statistiques, nous utilisons les valeurs critiques figurant dans le Tableau 4.15 de D'Agostino et Stephens (1986).

Notons que les estimateurs $\hat{\alpha}_X$ et $\hat{\beta}_X$ utilisés par Stephens sont les modifications des estimateurs des moments proposées par Cohen et Helm (1973) en rappelant que les estimateurs des moments des paramètres $\alpha$ et $\beta$ sont respectivement $(\bar{X} - S)$ et S où

$$\bar{X} = \sum_{i=1}^{n} X_i/n \text{ et } S = \left\{ \sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1) \right\}^{1/2} ;$$

voir (Johnson et al., 1994, pages 506-508). Rappelons aussi que les estimateurs du MLE sont respectivement $X_{(1)}$ et $(\bar{X} - X_{(1)})$.

Il est clair que le remplacement des estimateurs $\hat{\alpha}$ et $\hat{\beta}$ par d'autres estimateurs aura une influence sur la puissance des tests basés sur les statistiques de la f.r.e. Alors, nous proposons l'utilistation des deux estimateurs

$$\tilde{\alpha}_X = X_{(1)} - \tilde{\beta}_X/n \text{ et } \tilde{\beta}_X = S = \left\{ \sum_{i=1}^{n} (X_i - \bar{X})^2/(n-1) \right\}^{1/2} .$$

Les tests utilisant $\tilde{\alpha}_X$ et $\tilde{\beta}_X$ seront notés : $KSN, CMN, WAN$ et $ADN$.

Il est facile de vérifier que $\hat{\alpha}_X$ et $\tilde{\alpha}_X$ sont des estimateurs équivariants par translation et dispersion et que $\hat{\beta}_X$ et $\tilde{\beta}_X$ sont des estimateurs invariants par translation, mais équivariants par dispersion. Par conséquent, toutes les statistiques utilisées dans la présente section sont pivotales.

## 4.2 Tests basés sur la corrélation

Nous allons aussi examiner les statistiques fondées sur le coefficient de corrélation entre le vecteur $W_{(.)} = (W_{(1)}, ..., W_{(n)})$ et un vecteur de scores $c = (c_1, ..., c_n)$, où $c_1 \leq ... \leq c_n$, liée à $F_0$. Le coefficient de corrélation est défini comme

$$R(W_{(.)}, c) = \frac{\sum_{i=1}^{n} (W_{(i)} - \bar{W})(c_i - \bar{c})}{\left[\sum_{i=1}^{n} (W_{(i)} - \bar{W})^2 \sum_{i=1}^{n} (c_i - \bar{c})^2\right]^{1/2}} \tag{4}$$

et, en ce qui concerne les scores, nous allons nous concentrer sur le choix entre deux options, à savoir,

$$m_i = E\left(F_0^{-1}(U_{(i)})\right), i = 1, ..., n,$$

où $U_{(1)} < ... < U_{(n)}$ sont des statistiques d'ordre associées à un échantillon aléatoire de taille $n$ issu d'une distribution $U(0, 1)$, et

$$h_i = F_0^{-1}\left(E(U_{(i)})\right) = F_0^{-1}\left(\frac{i}{n+1}\right), i = 1, ..., n,$$

Et, comme les $W_i, 1 \leq i \leq n$, ne dépendent ni de $\alpha$ ni de $\beta$, il est clair que toute statistique définie à partir du coefficient de corrélation (4) est pivotale.

Nous considérons les deux statistiques basées sur le coefficient de corrélation :

$$Z(W_{(.)}, m) = n\{1 - R^2(W_{(.)}, m)\}$$

où $m_i = E\left\{-Log(1 - U_{(i)})\right\} = \sum_{j=1}^{i}(n - j + 1)^{-1}, 1 \leq i \leq n$, [voir Savage (1956)], sont les scores exacts, et

$$Z(W_{(.)}, h) = n\{1 - R^2(W_{(.)}, h)\}$$

où $h_i = -Log\left\{1 - E(U_{(i)})\right\} = -Log\left\{1 - i/(n + 1)\right\}, 1 \leq i \leq n$, sont les scores approximatifs. Ces deux tests ont été considérés par Stephens qui a évalué les valeurs critiques (D'Agostino et Stephens, 1986, Tables 5.6 and 5.7).

## 4.3 Tests basés sur différents estimateurs

### 4.3.1 Test de Shapiro-Wilk

Nous prenons la statistique de Shapiro-Wilk ($SW$), telle que définie par Shapiro et Wilk (1972) et étudiée par Metz et al. (1994). Cette statistique est :

$$SW = \frac{n(\bar{X} - X_{(1)})^2}{(n-1)S^2} \text{ où } S^2 = \frac{\sum_{i=1}^{n}(X_i - \bar{X})^2}{n-1}.$$

Pour les points critiques, nous allons utiliser les valeurs figurant dans le tableau 1 de Shapiro et Wilk (1972).

### 4.3.2 Test de Gini

La statistique du test de Gini telle que définie par Gail et Gastwirth (1978) est :

$$GI = \frac{\sum_{i=1}^{n-1} i(n-i)(X_{(i+1)} - X_{(i)})}{(n-1)\sum_{i=1}^{n} X_i}$$

Nous proposons aussi la statistique $GIM = 1/GI$ qui est une transformation monotone décroissante de la statistique de Gini.

### 4.3.3 Test de Kullback-Leibler

Pour ce test, nous utilisons la statistique de test normalisé

$$KL_{mn} = exp(-I_{mn}) = exp(H_{mn} - Log(\bar{W}) - 1)$$

basée sur l'estimateur de Vasicek $H_{mn}$ pour estimer la fonction de Kullback-Leibler de discrimination d'information entre deux distributions de données définie par

$$I(F : F_0) = \int_0^\infty f(x) Log\{\frac{f(x)}{f_0(x)}\}dx.$$

L'estimateur de Vasicek est donné par

$$H_{mn} = n^{-1} \sum_{i=1}^{n} Log\{\frac{n}{2m}(W_{(i+m)} - W_{(i-m)})\},$$

où la taille de la fenêtre $m$ est un nombre entier positif plus petit que $n/2$, $W_{(j)} = W_{(1)}$ si $j < 1, W_{(j)} = W_{(n)}$ si $j > n$. Ce test a été examiné par Ebrahimi et al. (1992) qui ont évalué les valeurs critiques (Ebrahimi et al., 1992, Tables 1 and 2).

## 4.4 Tests fondés sur des combinaisons de statistiques standardisées

Une fois l'étude de simulation sur la base des statistiques ci-dessus a été réalisée, nous avons remarqué qu'un groupe de tests MCR ont donné lieu à des puissances importantes pour un premier sous-ensemble d'alternatives, mais très faibles pour un deuxième sous-ensemble. De plus, aucun des tests n'a maintenu une puissance acceptable envers toutes les alternatives envisagées. Pour exploiter ce fait, nous allons examiner ici les tests basés sur le maximum de plusieurs des statistiques standardisées. En outre, grâce à l'utilisation de la technique des tests de MCR, nous serons en mesure de prendre en compte de la dépendance entre les statistiques des tests, par conséquent, éviter l'hypothèse de l'indépendance exigée souvent dans la littérature sur la combinaison de tests [voir Folks (1984) ou dans l'utilisation des approximations fondées sur les limites ou sur les arguments asymptotiques [voir Miller (1981), Westfall et Young (1993), Pesarin (2001) et Dufour et Khalaf (2002)].

La standardisation vise à assurer une comparabilité entre les statistiques et consiste simplement à soustraire la moyenne empirique de chaque statistique et à diviser le résultat par l'écart type empirique correspondant, où la moyenne empirique et l'écart type sont calculés à partir des valeurs observées et simulées des statistiques de test.

Formellement, si $V = (T_1, \ldots, T_k)'$ désigne un vecteur des $k$ statistiques sélectionnées, soit $V^{(0)} = (T_1^{(0)}, \ldots, T_k^{(0)})'$ le vecteur des statistiques obtenues de l'échantillon observé $X$ et soient $V^{(i)} = (T_1^{(i)}, \ldots, T_k^{(i)})'$, $i = 1, \ldots, N$, les vecteurs des statistiques obtenues des $N$ échantillons aléatoires simulés $X_1, \ldots, X_n$. Les statistiques sont standardisés comme suit :

$$\widetilde{T}_j^{(i)} = \frac{T_j^{(i)} - \overline{T}_j}{s_j}, \; j = 1, \ldots, \; k, \; i = 0, 1, \ldots, N,$$

où

$$\overline{T}_j = \frac{1}{N+1} \sum_{i=0}^{N} T_j^{(i)}, \quad s_j = \left\{ \frac{1}{N} \sum_{i=0}^{N} \left( T_j^{(i)} - \overline{T}_j \right)^2 \right\}^{1/2}, \; j = 1, \ldots, \; k\,.$$

Pour le vecteur des statistiques observées $V^{(0)}$ et pour chacun des vecteurs des statistiques simulées $\big(V^{(i)}, \; i = 1, \ldots, N\big)$, nous pouvons calculer les statistiques combinées suivantes :

$$\widehat{Q}\big(V^{(i)}\big) = \max_{1 \leq j \leq k} \big\{\widetilde{T}_j^{(i)}\big\}, \; i = 0, 1, \ldots, N, \tag{5}$$

et

$$\widehat{Q}_a\big(V^{(i)}\big) = \max_{1 \leq j \leq k} \big\{\big|\widetilde{T}_j^{(i)}\big|\big\}, \; i = 0, 1, \ldots, N. \tag{6}$$

Le test combiné basé sur la statistique $\widehat{Q}$ rejette l'hypothèse nulle quand le maximum des statistiques standardisées est grand, tandis que celui basé sur $\widehat{Q}_a$ la rejette lorsque le maximum des valeurs absolues des statistiques standardisées est grand. Dans (5) et (6), $\widehat{Q}\big(V^{(0)}\big)$ et $\widehat{Q}_a\big(V^{(0)}\big)$ représentent les statistiques associées à l'échantillon observé, alors que $\widehat{Q}\big(V^{(i)}\big)$ et $\widehat{Q}_a\big(V^{(i)}\big)$ pour $i \neq 0$ peuvent être interprétées comme des valeurs fondées sur la simulation des échantillons.

Il est facile de voir que les variables $\widehat{Q}\big(V^{(i)}\big)\,,\ i = 0,\, 1,\, \ldots\,,\, N,$ sont interchangeables sous $H_0$ et de même pour $\widehat{Q}_a\big(V^{(i)}\big)\,,\ i = 0,\, 1,\, \ldots\,,\, N$. En conséquence, on peut écrire :

$$\widehat{Q}\big(V^{(i)}\big)\,,\ i = 0,\, 1,\, \ldots\,,\, N, \ \text{ sont interchangeables sous } H_0\,,$$

et de même pour $\widehat{Q}_a\big(V^{(i)}\big)\,,\ i = 0,\, 1,\, \ldots\,,\, N$. En conséquence, on peut écrire :

$$\mathsf{P}\big[\hat{p}_N\big(\widehat{Q}^{(0)}\big) \le \alpha\big] \le \mathsf{P}\big[\tilde{p}_N\big(\widehat{Q}^{(0)}\big) \le \alpha \,|\, H_0\big] = \frac{I[\alpha(N+1)]}{N+1},\ \ 0 \le \alpha \le 1,$$

où $\widehat{Q}^{(i)} \equiv \widehat{Q}\big(V^{(i)}\big)\,,\ i = 0,\, 1,\, \ldots\,,\, N$, et $\hat{p}_N, \tilde{p}_N$ sont définis, respectivement, comme dans (1) et (3) avec $T_i$ remplacé par $\widehat{Q}^{(i)}$. Bien sûr, il en va de même pour les tests basés sur $\widehat{Q}_a^{(i)} \equiv \widehat{Q}_a\big(V^{(i)}\big),\ i = 0,\, 1,\, \ldots\,,\, N.$

Dans notre étude de simulations ci-dessous, aucun des tests individuels présentés n'a maintenu la meilleure puissance envers toutes les alternatives considérées. Ceci nous a incité à construire des statistiques de test combinées. Ci-dessous, nous examinerons trois de ces cas des statistiques de test combinées :

$$\widehat{Q}_i \equiv \widehat{Q}(V_i)\,,\ i = 1,\, \ldots\,,\, 3\,,$$

où

$$V_1 = (ADN,\, KL)'\,, \quad V_2 = (GI,\, SW)' \ \text{ et } \ V_3 = (ADN, KL, GI, SW)'.$$

Le premier choix $(V_1)$ met l'accent sur deux mesures de distance entre les deux distributions empiriques $ADN$ et $KL$, tandis que le second $(V_2)$ utilise également des différences et des rapports de mesures de dispersion basées sur les moments et les statistique d'ordre des deux distributions, et devraient donc fournir plus de sensibilité à la différence qui influe sur les moments et les statistiques d'ordre. Le vecteur $V_3$ a regroupé les quatre composantes des vecteurs $V_1$ et $V_2$ dans le but de voir l'effet de plusieurs statistiques de natures différentes.

# 5 Étude de simulation

Dans cette section, nous présentons des comparaisons de performance (du point de vue niveau et puissance) entre les différents tests étudiés. Les tests de MCR sont basés sur N = 99 échantillons aléatoires indépendants identiquement distribués simulés sous l'hypothèse nulle $H_0$ (c'est à dire des échantillons de la loi $E(0,1)$). Afin de mesurer les niveaux et les puissances des tests étudiés, nous avons utilisé 10 000 réitérations des différents tests sous chaque spécification considérée. Le niveau des tests est de $5\%$. Le taux de rejet de l'hypothèse $H_0$ sur les 10 000 réitérations constitue le niveau expérimental du test si l'échantillon est généré de la loi $E(0,1)$ et la puissance sinon. Afin d'évaluer les puissances empiriques des tests , nous avons généré des échantillons provenant des distributions suivantes [1] : Valeur absolue d'une loi normale standard $|N|$, la loi béta B(1,1), la loi Gamma $\Gamma$(2), la loi lognormal ($\Lambda(4,\ 1.5^2)$), la loi student(5) en valeur absolue $|t(5)|$ et la loi uniforme continue $U[0,1]$.

| | Tests Originaux | | | | Tests MCR | | | |
|---|---|---|---|---|---|---|---|---|
| n | 10 | 15 | 25 | 50 | 10 | 15 | 25 | 50 |
| KS | 4.6 | 5.1 | 5.1 | 4.8 | 4.9 | 5.1 | 5.1 | 4.9 |
| CM | 5.1 | 5.4 | 5.5 | 5.3 | 5.0 | 5.1 | 5.0 | 5.0 |
| W | 5.3 | 5.4 | 5.3 | 5.2 | 4.9 | 5.1 | 5.1 | 4.9 |
| AD | 5.2 | 5.3 | 5.3 | 5.0 | 5.0 | 5.1 | 5.1 | 5.0 |
| $Z(W_{(.)}, m)$ | 5.0 | 5.3 | 5.7 | 5.0 | 5.0 | 5.1 | 5.0 | 5.0 |
| $Z(W_{(.)}, h)$ | 4.8 | 5.2 | 5.3 | 5.0 | 5.1 | 5.0 | 5.0 | 5.1 |
| SW | 5.4 | 5.0 | 5.3 | 4.8 | 5.0 | 5.0 | 5.1 | 5.0 |
| GI | 10.1 | 10. | 6.5 | ND | 5.0 | 5.1 | 5.0 | 5.0 |
| KL | 5.7 | 5.0 | 4.8 | 3.8 | 5.0 | 5.1 | 5.1 | 5.0 |

TAB. 1 – *Niveaux Empiriques pour les tests d'ajustement à la distribution exponentielle* $\alpha = 0.05$.

Pour tous nos calculs et simulations, nous avons constitué un programme en Fortran 90 avec l'extension de la programmatique *International Mathematical and Statistical Libraries* (IMSL). Ce programme a été compilé et exécuté à l'aide du système d'exploitation Unix.

Les résultats du tableau 1 montrent que la majorité des tests originaux basés sur des valeurs critiques ont souvent un problème de niveaux tandis que les tests de MCR ont des niveaux exacts.

L'analyse des résultats des tableaux 2, 3 et 4 montrent que :

**a)** En utilisant les estimateurs $\hat{\alpha}$ et $\hat{\beta}$, nous pouvons conclure :

- Les quatre tests basés sur la f.r.e. ont des puissances semblables, mais le test $CM$ est le plus performant,

- Le test fondé sur la transformation $GIM$ de la statistique $GI$ et le test $SW$ sont souvent les plus puissants envers les cinq alternatives : $|N|, B, \Gamma, |t|$ et $U$. Par contre, leurs puissances s'annulent si l'échantillon provient d'une loi $\Lambda$,

- Le tes $KL$ présente la meilleure puissance pour les alternatives $B$ et $U$ et des puissances modérées ou faibles pour les autres distributions,

- Les deux tests $Z(W_{(.)}, m)$ et $Z(W_{(.)}, h)$ basés sur la corrélation ne présentent aucune performance. Nous signalons qu'à l'exception de la distribution $\Lambda$, le test $Z(W_{(.)}, m)$ présente une puissance supérieure à celle du $Z(W_{(.)}, h)$,

**b)** Il est clair que le remplacement des estimateurs $\hat{\alpha}$ et $\hat{\beta}$ par $\tilde{\alpha}_X$ et $\tilde{\beta}_X$ améliore sensiblement les puissances des quatre tests $KSN, CMN, WAN$ et $ADN$ fondés sur la f.r.e.

**c)** Les tests $\widehat{Q}_1, \widehat{Q}_2$ et $\widehat{Q}_3$ fondés sur les combinaisons maintiennt des bonnes puissances robustes envers toutes les alternatives.

1. Les définitions précises de ces lois sont présentées dans la majorité des manuels de statistique et probabilités.

| | Tests utilisant $(\hat{\alpha}, \hat{\beta})$ | | | | | | |
|---|---|---|---|---|---|---|---|
| Test / F | E | $\lvert N\rvert$ | B | $\Gamma$ | U | $\Lambda$ | $\lvert t\rvert$ |
| KS | 4.9 | 9.0 | 23.8 | 9.4 | 23.9 | 30.2 | 6.6 |
| CM | 4.9 | 9.7 | 30.3 | 10.3 | 30.3 | 33.0 | 6.8 |
| W | 4.9 | 8.7 | 27.3 | 9.3 | 27.3 | 28.2 | 6.4 |
| AD | 5.0 | 8.5 | 28.9 | 8.9 | 28.8 | 35.8 | 6.2 |
| $Z(W_{(.)}, m)$ | 4.9 | 9.1 | 40.9 | 7.6 | 41.1 | 19.9 | 6.7 |
| $Z(W_{(.)}, h)$ | 4.9 | 5.0 | 26.9 | 4.9 | 27.1 | 27.2 | 5.5 |
| SW | 5.1 | 15.5 | 45.1 | 15.4 | 45.0 | 0.6 | 9.3 |
| GI | 5.0 | 8.9 | 30.8 | 9.3 | 30.8 | 36.9 | 6.5 |
| GIM | 5.1 | 15.4 | 43.4 | 15.6 | 43.3 | 0.6 | 9.3 |
| KL | 5.0 | 13.2 | 49.8 | 12.1 | 50.0 | 13.1 | 8.0 |
| | Tests utilisant $(\tilde{\alpha}, \tilde{\beta})$ | | | | | | |
| KSN | 5.0 | 13.4 | 37.8 | 13.3 | 37.7 | 21.5 | 8.5 |
| CMN | 5.0 | 14.3 | 40.8 | 14.4 | 40.7 | 20.9 | 8.9 |
| WN | 5.0 | 11.1 | 34.7 | 11.0 | 34.5 | 27.9 | 7.6 |
| ADN | 5.0 | 14.5 | 43.4 | 14.4 | 43.4 | 20.1 | 8.9 |
| | Tests MCR Combinés | | | | | | |
| $\widehat{Q}_1$ | 5.0 | 14.1 | 47.0 | 13.5 | 46.9 | 18.8 | 8.6 |
| $\widehat{Q}_2$ | 5.0 | 11.7 | 37.7 | 11.7 | 37.6 | 31.5 | 7.6 |
| $\widehat{Q}_3$ | 5.0 | 12.6 | 44.0 | 12.2 | 44.0 | 27.3 | 8.0 |

TAB. 2 – *Niveaux et puissances empiriques pour les tests MCR d'ajustement à la distribution exponentielle* $n = 10$ *et* $\alpha = 0.05$

| | Tests utilisant $(\hat{\alpha}, \hat{\beta})$ | | | | | | |
|---|---|---|---|---|---|---|---|
| Test / F | E | $\|N\|$ | B | $\Gamma$ | U | $\Lambda$ | $\|t\|$ |
| KS | 5.0 | 18.9 | 59.1 | 24.9 | 59.2 | 65.4 | 9.8 |
| CM | 5.0 | 22.3 | 74.3 | 29.3 | 74.1 | 70.4 | 10.9 |
| W | 5.0 | 18.3 | 67.5 | 24.8 | 67.6 | 57.7 | 9.9 |
| AD | 5.0 | 20.3 | 74.6 | 27.4 | 74.5 | 72.1 | 9.8 |
| $Z(W_{(.)}, m)$ | 5.1 | 10.9 | 86.5 | 7.6 | 86.7 | 48.2 | 8.0 |
| $Z(W_{(.)}, h)$ | 5.0 | 3.3 | 60.7 | 3.3 | 61.1 | 53.2 | 6.6 |
| SW | 5.0 | 36.8 | 91.5 | 37.7 | 91.4 | 0.0 | 15.1 |
| GI | 5.1 | 23.9 | 78.9 | 29.2 | 78.8 | 74.4 | 10.3 |
| GIM | 5.0 | 35.5 | 87.3 | 41.3 | 87.2 | 0.0 | 15.9 |
| KL | 5.0 | 24.7 | 95.8 | 23.6 | 95.9 | 39.1 | 9.5 |
| | Tests utilisant $(\tilde{\alpha}, \tilde{\beta})$ | | | | | | |
| KSN | 5.1 | 29.2 | 83.1 | 32.0 | 83.0 | 63.1 | 13.1 |
| CMN | 5.1 | 31.6 | 86.3 | 35.2 | 86.3 | 64.0 | 13.9 |
| WN | 5.1 | 24.9 | 80.6 | 26.7 | 80.6 | 64.8 | 11.6 |
| ADN | 5.1 | 32.0 | 88.8 | 34.9 | 88.8 | 64.9 | 13.8 |
| | Tests MCR Combinés | | | | | | |
| $\widehat{Q}_1$ | 5.0 | 29.9 | 93.6 | 31.5 | 93.6 | 62.6 | 12.4 |
| $\widehat{Q}_2$ | 5.0 | 28.8 | 87.0 | 31.6 | 86.9 | 70.7 | 11.9 |
| $\widehat{Q}_3$ | 5.0 | 28.6 | 92.9 | 30.4 | 92.9 | 67.4 | 11.8 |

TAB. 3 – *Niveaux et puissances empiriques pour les tests MCR d'ajustement à la distribution exponentielle* $n = 25$ *et* $\alpha = 0.05$

## 6 Conclusion

Cet article décrit des stratégies de tests de MCR dont la validité peut être démontrée pour une grande classe de modèle d'événements statistiques. De plus, la conclusion commune à toute cette étude est que les tests basés sur des valeurs critiques ont souvent un problème de niveau pour des échantillons relativement faible (<50). Nous pouvons souligner que les tests de MCR fondés sur des statistiques combinées que nous proposons sont les plus puissants de tous les tests. Par conséquent, des raffinements et des extensions selon les orientations évoquées plus haut, permettent d'anticiper des résultats fort utiles en statistique et économétrie du point de vue théorique et empirique.

| | Tests utilisant $(\hat{\alpha}, \hat{\beta})$ | | | | | | |
|---|---|---|---|---|---|---|---|
| Test / F | E | $\|N\|$ | B | $\Gamma$ | U | $\Lambda$ | $\|t\|$ |
| KS | 5.0 | 36.5 | 91.0 | 54.5 | 90.9 | 90.7 | 15.3 |
| CM | 5.0 | 44.6 | 97.7 | 62.9 | 97.7 | 93.6 | 17.6 |
| W | 5.0 | 35.0 | 95.3 | 54.4 | 95.3 | 85.3 | 15.8 |
| AD | 5.0 | 42.2 | 98.2 | 62.8 | 98.2 | 93.9 | 16.3 |
| $Z(W_{(.)}, m)$ | 4.9 | 16.7 | 99.7 | 8.2 | 99.7 | 73.2 | 9.5 |
| $Z(W_{(.)}, h)$ | 5.0 | 3.9 | 94.4 | 2.9 | 94.3 | 75.2 | 8.4 |
| SW | 5.1 | 67.0 | 99.9 | 67.1 | 99.9 | 0.0 | 22.0 |
| GI | 5.0 | 50.4 | 98.7 | 64.2 | 98.6 | 95.2 | 16.8 |
| GIM | 5.1 | 63.6 | 99.4 | 75.3 | 99.5 | 0.0 | 25.3 |
| KL | 4.9 | 39.1 | 100. | 42.2 | 100. | 76.7 | 10.1 |
| | Tests utilisant $(\tilde{\alpha}, \tilde{\beta})$ | | | | | | |
| KSN | 5.0 | 54.9 | 99.2 | 61.9 | 99.2 | 91.7 | 19.2 |
| CMN | 5.0 | 58.9 | 99.5 | 65.9 | 99.5 | 92.3 | 20.1 |
| WN | 5.0 | 48.1 | 98.9 | 54.7 | 98.9 | 91.6 | 17.3 |
| ADN | 5.0 | 59.2 | 99.7 | 65.9 | 99.7 | 92.7 | 19.8 |
| | Tests MCR Combinés | | | | | | |
| $\widehat{Q}_1$ | 5.0 | 55.5 | 100. | 61.5 | 100. | 91.8 | 17.3 |
| $\widehat{Q}_2$ | 5.1 | 56.7 | 99.6 | 63.7 | 99.7 | 94.3 | 18.2 |
| $\widehat{Q}_3$ | 5.0 | 54.6 | 100. | 61.2 | 100. | 93.1 | 16.9 |

TAB. 4 – *Niveaux et puissances empiriques pour les tests MCR d'ajustement à la distribution exponentielle* $n = 50$ *et* $\alpha = 0.05$

## Références

Anderson, T. W. et D. A. Darling (1954). A test of goodness-of-fit. *Journal of the American Statistical Association 49*, 765–769.

Barnard, G. A. (1963). Comment on 'The spectral analysis of point processes' by M. S. Bartlett. *Journal of the Royal Statistical Society, Series B 25*, 294.

Birnbaum, Z. W. (1974). Computers and unconventional test-statistics. In F. Proschan et R. J. Serfling (Eds.), *Reliability and Biometry*, pp. 441–458. Philadelphia, PA : SIAM.

Cohen, A. C. et F. R. Helm (1973). Estimator in the exponential distribution. *Technometrics 15*, 415–418.

Cramér, H. (1928). On the composition of elementary errors. *Skandinavisk Aktuarietidskrift 11*, 141–180.

D'Agostino, R. B. et M. A. Stephens (1986). *Goodness-of-Fit Techniques*. New York : Marcel Dekker.

Dufour, J.-M. (2006). Monte Carlo tests with nuisance parameters : A general approach to finite-sample inference and nonstandard asymptotics in econometrics. *Journal of Econometrics 133*, 443–477.

Dufour, J.-M. et A. Farhat (2002). Exact nonparametric two-sample homogeneity tests. In C. Huber-Carol, N. Balakhrishnan, M. Nikulin, et M. Mesbah (Eds.), *Proceedings of the 2000 International Workshop on "Goodness-of-fit Tests and Validity of Models"*, Chapter 33, pp. 435–448. Boston, Massachusetts : Birkhaüser.

Dufour, J.-M., A. Farhat, L. Gardiol, et L. Khalaf (1998). Simulation-based finite sample normality tests in linear regressions. *The Econometrics Journal 1*, 154–173.

Dufour, J.-M., A. Farhat, et L. Khalaf (2004). Tests multiples simulés et tests de normalité basés sur plusieurs moments dans les modèles de régression. *L'Actualité économique 80*, 593–618.

Dufour, J.-M. et L. Khalaf (2002). Exact tests for contemporaneous correlation of disturbances in seemingly unrelated regressions. *Journal of Econometrics 106*(1), 143–170.

Dwass, M. (1957). Modified randomization tests for nonparametric hypotheses. *Annals of Mathematical Statistics 28*, 181–187.

Ebrahimi, N., M. Habibullah, et E. S. Soofi (1992). Testing exponentiality based on Kullback-Leibler information. *Journal of the Royal Statistical Society, Series B 54*(3), 739–748.

Folks, J. L. (1984). Combination of independent tests. In P. R. Krishnaiah et P. K. Sen (Eds.), *Handbook of Statistics 4 : Nonparametric Methods*, pp. 113–121. Amsterdam : North-Holland.

Gail, M. H. et J. L. Gastwirth (1978). A scale-free goddness-of-fit test for the exponential distribution based on the Lorenz curve (in theory and methods). *Journal of the American Statistical Association 73*, 787–793.

Johnson, N. L., S. Kotz, et N. Balakrishnan (1994). *Continuous Univariate Distributions, Volume 1* (Second ed.). New York : John Wiley & Sons.

Kolmogorov, A. N. (1933). Sulla determinazione empirica di una legge di distribizione. *Giornale dell' Instituto Italiano degli Atturai 4*, 83–91.

Lilliefors, W. H. (1969). On the Kolmogorov-Smirnov test for the exponential distribution with mean unknown. *Journal of the American Statistical Association 64*, 387–389.

Maurer, W. et B. H. Margolin (1976). Tests of Kolmogorov-Smirnov type for exponential data with unknown scale, and related problems. *Biometrika 63*, 149–160.

Metz, J. A. J., P. Haccou, et E. Meelis (1994). On the Shapiro-Wilk test and Darling's test for exponentiality. *Biometrics 50*, 527–530.

Miller, Jr., R. G. (1981). *Simultaneous Statistical Inference* (Second ed.). New York : Springer-Verlag.

Pesarin, F. (2001). *Multivariate Permutation Tests with Applications in Biostatistics*. New York : John Wiley & Sons.

Savage, I. R. (1956). Contributions to the theory of rank order statistics : Two-sample case. *Annals of Mathematical Statistics 27*, 590–615.

Shapiro, S. S. et M. B. Wilk (1972). An analysis of variance test for the exponential distribution (complete samples). *Technometrics 14*, 355–370.

Smirnov, N. V. (1939). Sur les écarts de la courbe de distribution empirique (Russian/French summary). *Matematičeskiĭ Sbornik N.S. 6*, 3–26.

Stephens, M. A. (1974). EDF statistics for goodness-of-fit and some comparisons. *Journal of the American Statistical Association 69*, 730–737.

Watson, G. S. (1961). Goodness-of-fit tests on a circle. *Biometrika 48*, 109–114.

Westfall, P. H. et S. S. Young (1993). *Resampling-Based Multiple Testing : Examples and Methods for $p$-Value Adjustment*. New York : John Wiley & Sons.

## Summary

The exponential distribution is widely used for modeling duration data in statistics, econometrics and finance. Thus testing for exponentiality is an important problem when studying such data. Most of the tests proposed in the literature focus on exponential distributions with only one parameter, and the available tables only consider a limited number of sample sizes and levels, which can be the source of size and power problems. In this paper, we first propose to use the technique of Monte Carlo tests in order to control the size of various exponentiality tests involving both a scale and a location nuisance parameter. We show both theoretically and by simulation that the tests obtained in this way are exact for any given sample size. Second, we introduce modified procedures based on simple moment estimators and show that the latter yield power improvements over ealier alternative tests. Third, we extend some tests proposed for exponential distributions with one parameter, such as tests based on Gini coefficients, to exponential distributions with two parameters. Fourth, we propose and show how to obtain exact exponentiality tests based on kernel-type density estimators, which can have better power than tests based on empirical distribution functions and the Shapiro-Wilk test. Fifth, we suggest test procedures combining several tests and show these can also yield power improvements. Finally, the procedures considered are compared in a Monte Carlo experiment.