

République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

UNIVERSITE MOHAMED KHIDER BISKRA
FACULTE DES SCIENCES ET DES SCIENCES DE L'INGENIEUR
DEPARTEMENT DE MATHEMATIQUES

MEMOIRE

Présenté en vue de l'obtention du diplôme de

Magister en Mathématiques

Par

BERKANE Hassiba

Thème

Méthodes du Bootstrap pour les queues de distributions

Option

Analyse & Modèles Aléatoires

Soutenu publiquement le : 17/05/2005

Devant le jury :

Président :	B. MEZERDI	PR.	U.M.K. Biskra.
Rapporteur :	A. NECIR	M.C.	U.M.K. Biskra.
Examineur :	K. BOUKHETALA	PR.	U.S.T.H.B. Alger.
Examineur :	S. BAHLALI	DR. M.A.C.	U.M.K. Biskra.

TABLE DES MATIÈRES

0.1. Notations	3
0.2. Introduction	5
CHAPITRE 1. NOTIONS DE STATISTIQUE	7
1.1. Théorie de l'estimation	7
1.1.1. Modèles Statistiques	7
1.1.2. Estimateur	8
1.2. Méthodes paramétriques	9
1.2.1. Biais et risque	9
1.2.2. Intervalles de confiance	12
1.2.3. Test d'hypothèses	14
1.3. Méthodes non-paramétriques	17
1.3.1. Estimation non-paramétrique	17
1.3.2. Modèle d'échantillonnage d'une loi sur $\mathbb{R}$	17
1.3.3. Tests non-paramétriques	20
CHAPITRE 2. PROCEDURE DU BOOTSTRAP	23
2.1. Principe du Bootstrap	23
2.2. Validité du Bootstrap	28
2.2.1. Théorie asymptotique du bootstrap	28
2.2.2. Théorème centrale limite	30
2.3. Consistance du bootstrap	33
2.3.1. Distribution de la Statistique	37

CHAPITRE 3. UTILISATION DU BOOTSTRAP POUR LES PROBLEMES STATISTIQUES	43
3.1. Méthodes de rééchantillonnage	43
3.1.1. Bootstrap des individus	43
3.2. Erreur standard et biais d'un paramètre	44
3.2.1. Estimation de l'erreur-standard	44
3.2.2. Estimation du biais	48
3.2.3. Estimations par le jackknife	50
3.3. Bootstrap pour les tests d'hypothèse	53
3.4. Intervalle de Confiance et Bootstrap	56
3.4.1. Méthode de l'erreur-standard	56
3.4.2. Méthode des pourcentiles simples	57
3.4.3. Méthode des pourcentiles corrigés pour le biais	58
3.4.4. Méthode des pourcentiles avec correction pour le biais et accélération	58
3.4.5. Méthode du bootstrap-t	59
3.5. Techniques du bootstrap pour les modèles de régression	63
3.5.1. Bootstrap des Résidus	63
3.5.2. Les intervalles de prédiction bootstrap	65
3.5.3. Bootstrap par paires	68
CHAPITRE 4. BOOTSTRAP POUR LES VALEURS EXTRÊMES	72
4.1. Méthodes classiques d'estimation des quantiles extrêmes	72
4.1.1. Méthode des valeurs extrêmes	73
4.1.2. Méthode des excès	74
4.1.3. Méthode des quantiles	76
4.2. Test d'adéquation pour les queues de distributions	79
4.2.1. Test d'adéquation via la distance L^2 -Wasserstein pondérés	80
4.3. Simulations	89

RESUME

Ce mémoire présente une étude théorique sur les différentes méthodes du Bootstrap et l'utilisation de cette technique de ré échantillonnage dans la statistique d'inférence pour le calcul de l'erreur standard, du biais d'un estimateur et la détermination de l'intervalle de confiance d'un paramètre estimé. Nous appliquons ces méthodes dans les testes des modèles de régression et les modèle de type Pareto, ce qui donne des meilleurs approximations. Du coté numérique, notre étude est fonder sur le logiciel statistique R version 2.0.1.

Mots- clés: Bootstrap, erreur standard, biais, jackknife, moyenne, médiane, variance, quantile, intervalle de confiance, régression, valeur extrême, queue indexé, queue lourde, bon ajustement, distance de Wasserstein.

ABSTRACT

This memory presents a theatrical study on various methods of Bootstrap and the uses of this technique to resampling statistics of inference for computing the standard error, bias of an estimator and the determination of the confidence interval of an estimated parameter. We apply these methods to testing models of regression and the Pareto type distribution, which gives better approximations. The numerical study is based on the statistical software R version 2.0.1.

Key words: Bootstrap, standard error, bias, jackknife, mean, median, variance, quantile, confidence intervals, regression, extreme value, tail index, heavy tails, goodness of fit, Wasserstein distance.

ملخص

قمنا في هذه المذكرة بعرض الدراسة النظرية لمختلف طرق العينة المتكررة، إذ تطرقنا لمختلف استعمالاتها في الإحصاء الاستدلالي و ذلك لإيجاد الخطأ المعياري، الانحراف، مجالات الثقة و في الانحدارات. كما طبقنا طرق العينة المتكررة على أذنان التوزيعات من أجل نموذج باريتو لإيجاد أحسن التقريبات لاختبار عينة كبيرة الحجم ذات متغيرات مستقلة.

الكلمات الجوهرية: الخطأ المعياري، الانحراف، مجالات الثقة، المعدل، المتوسط، التباين، الانحدار، القيمة الحدية، دليل أذنان التوزيعات، أذنان كبيرة.