Hacettepe Journal of Mathematics and Statistics
Volume 45 (2) (2016), 561 – 573

Modified Welch test statistic for ANOVA under Weibull distribution

Derya Karagöz*†‡

Abstract

A modification to Welch test statistic is proposed to test the equality of population means of various groups under a Weibull distribution. The proposed test statistic is simple and corresponds to the standard Welch test statistic in which the maximum likelihood mean and variance estimators are replaced with robust estimators based on quantile, quantile least square and repeated median. The influence function and breakdown point of these robust estimators are obtained to show their robustness properties. In the simulation study, various experimental designs are considered to evaluate the performance of proposed modified Welch classical ANOVA tests in terms of the type I-errors studies via simulation study.

Keywords: ANOVA, Welch Test Statistic, Robustness of estimators.

2000 AMS Classification: 62K25, 62J10, 62E17

Received : 03.09.2014 *Accepted* : 10.03.2015 *Doi* : 10.15672/HJMS.2015569632

1. Introduction

Analysis of variance (ANOVA) is one of the most used model which can be seen in many fields such as medicine, engineering, agriculture, education, psychology, sociology and biology to investigate the source of the variations. In general, the main interest is in testing the homogeneity of group means using the classical ANOVA which uses F-test statistic. One-way ANOVA is based on assumptions that the normality of the observations and the homogeneity of group variances. If the assumptions of normality and homogeneity of variances are invalid and also outliers are present, classical ANOVA does not give accurate results. Therefore, test statistics based on robust methods should be used instead of the classical ANOVA.

*Department of Statistics, Hacettepe University, Beytepe, Ankara, Turkey, Email: deryacal@hacettepe.edu.tr
†Corresponding Author.
‡This work forms part of the author's PhD thesis [8]. The author wishes to thank her research supervisor Prof. Dr. Tülay Saraçbaşı for her advice and support throughout.

The one-way ANOVA under the violation of assumptions has been studied extensively. To deal with non-normal data and/or heteroscedastic variances across groups, many alternatives such as Q, Welch, Brown-Forsythe and Modified Brown-Forsythe tests have been developed instead of classical ANOVA. The statistic Q has been extensively studied by many authors under a variety of assumptions. It is one of the most commonly used test statistic for homogeneity at population means in meta-analysis, see for example [5], [12]. [3] showed that under the null hypothesis Q asymptotically follows a Chi-Square distribution. [7] and [13] derived improved approximations to the distribution of Q under the null hypothesis; these approximations are more accurate for small sample sizes of groups. [9] extended the methods of Welch to find approximate distributions to Q under alternative hypotheses. [9] provide approximations for the non-null distributions of their weighted statistics which are found to be useful in obtaining approximations to the power of the Welch test. A number of authors have discussed extensions of the Welch methods based on the use of robust estimators for the population location and scale parameters. Notable among these are the efforts of [14], [15], and the references contained there in. [10] consider three common robust estimators: Huber's proposal two estimator of location and scale, Hampel's M-estimator of location with scale estimated by the median absolute deviation (MAD), and the trimmed mean with scale estimated by the Winsorized standard deviation.

One of the important assumptions of ANOVA is normality. However, in the application this assumption does not work for the real life data modeled by the exponential, Weibull or lognormal distributions especially in reliability, engineering and life science field. The characteristics of these distributions can be explained by Weibull distribution which is also known as Extreme Value Type III minimum distribution. This has made it extremely popular among reliability engineering, biology and medicine. This distribution is the most commonly used distribution for modeling reliability data, because it represents a wide range of asymmetric distributions. Moreover, ANOVA cannot handle censored or interval data because of the non-normality. The simplest possible lifetime distribution is exponential distribution. However, its constant hazard rate is improper and unrealistic in many cases. Gamma distribution is another candidate distribution for lifetimes. Nevertheless, distribution function or survival function of gamma distribution cannot be expressed in a closed form if the shape parameter is not an integer. Since it is in terms of an incomplete gamma function, one needs to obtain the distribution function, survival function or the hazard rate by numerical integration. This makes gamma distribution little bit unpopular compared to the Weibull distribution, which has a nice distribution function, survival function and hazard function [6]. The Weibull distribution was introduced by the Swedish physicist Weibull (1951). He claimed that his distribution applied to a wide range of problems and illustrated this point with seven examples ranging from the strength of steel to the height of adult males in the British Isles [1]. It has been used in many different fields like material science, engineering, physics, chemistry, meteorology,medicine, pharmacy, quality control, biology, geology, geography, economics and business.

This paper proposes a modified Welch test statistic to test the equality of population means of groups by utilizing robust estimators for the means and variances of Weibull distribution with outlier, and evaluates the performance of modified test in terms of the type I-errors via simulation study. The modified test statistic is called robust Welch (RW) test statistic. Since it is obtained by using robust mean and variance estimators based on quantile (Q), quantile least square (QLS) and repeated median ($Rmed$) instead of maximum likelihood. The influence function (IF) and breakdown point (BP) of robust estimators of mean and variance are obtained to show their robustness properties. The behavior of the developed robust test statistic is examined by Monte-Carlo simulation study. In the simulation study, various experimental designs are considered such as balanced and unbalanced sample sizes for k=3,6 groups with homogeneous and heterogeneous variances. The type I errors of the improved robust test statistic and classical ANOVA under the Weibull distribution are obtained.

The remainder of the paper is organized as follows. Section 2 introduces robust modified Welch test statistics. Section 3 gives explicit robust estimators of the mean and variance of Weibull distribution. The most important robustness measures are IF and BP that are derived in Section 4. To show the performance of the considered test statistic, a simulation study and the results are presented in Section 5. The last section summarizes the conclusions of the study.

2. Robust Welch Test Statistic

The Welch test statistic was proposed by [13] as following:

$$W = \frac{\hat{q}_w}{k-1}\{1 + \frac{2(k-2)\hat{A}}{k^2-1}\}^{-1}. \tag{2.1}$$

where

$$\hat{A} = \sum_{i=1}^{k}\left[1 - (\hat{w}_i/\sum_{j=1}^{k}\hat{w}_j)\right]^2/v_i \tag{2.2}$$

$$q_w \equiv \sum_{i=1}^{k}\hat{w}_i(\hat{\mu}_i - \hat{\mu}_w)^2 \tag{2.3}$$

and $v_i = n_i - 1$ is the degrees of freedom for i. sample. In (2.2) and (2.3), $\hat{\mu}_i$ is maximum likelihood estimator of the mean for each sample, $\hat{\sigma}_i^2$ is the maximum likelihood estimator of variance and $\hat{w}_i \equiv n_i/\hat{\sigma}_i^2$ is the estimator of weights based on variance estimator. If the appropriate weights are known, the value of estimation is

$$\hat{\mu}_w = \sum_{i=1}^{k}\hat{w}_i\hat{\mu}_i/\sum_{i=1}^{k}\hat{w}_i. \tag{2.4}$$

The Welch test statistic has approximately F_{k-1,v_w} distribution with $k-1$, $v_w = \frac{(k^2-1)}{3\hat{A}}$ degrees of freedom [13].

In this study, a modification to the Welch test statistic is proposed under a Weibull distribution. The test statistic is simple and corresponds to the standard Welch test statistic in which the maximum likelihood mean and variance estimators are replaced with robust estimators based on Q, QLS and $Rmed$. So the robust Welch test statistic is given by

$$RW = \frac{\hat{q}_{rw}}{k-1}\{1 + \frac{2(k-2)\hat{A}_r}{k^2-1}\}^{-1}. \tag{2.5}$$

where

$$(2.6) \qquad \hat{A}_r = \sum_{i=1}^{k} \left[1 - (\hat{w}_{ri} / \sum_{j=1}^{k} \hat{w}_{rj}) \right]^2 / v_i$$

$$(2.7) \qquad q_{rw} \equiv \sum_{i=1}^{k} \hat{w}_{ri} (\hat{\mu}_{ri} - \hat{\mu}_{rw})^2$$

and $v_i = n_i - 1$ is the degrees of freedom for i. sample.

In (2.7), $\hat{\mu}_{ri}$ is robust estimator of mean for each sample, $\hat{\sigma}_{ri}^2$ is robust estimator of variance and $\hat{w}_{ri} \equiv n_i / \hat{\sigma}_{ri}^2$ is the robust estimator of weights based on variance estimator. If the appropriate weights are known, the value of robust estimation are

$$(2.8) \qquad \hat{\mu}_{rw} = \sum_{i=1}^{k} \hat{w}_{ri} \hat{\mu}_{ri} / \sum_{i=1}^{k} \hat{w}_{ri}.$$

In (2.7) and (2.8), $\hat{\mu}_{ri}$ and $\hat{\sigma}_{ri}^2$ are the robust Q, QLS and $Rmed$ estimators of mean and variance for Weibull distribution.

The robust Welch test statistic has approximately $F_{k-1, v_{rw}}$ distribution with $k - 1$, $v_{rw} = \frac{(k^2-1)}{3\hat{A}_r}$ degrees of freedom.

3. Robust Estimators of Weibull Distribution

The estimations of mean and variance of Weibull distribution are a basic subject of the paper. The density function $f(x; \lambda, \beta) = \frac{\beta}{\lambda} (x/\lambda)^{\beta-1} \exp[-(x/\lambda)^{\beta}]$, $x, \lambda, \beta > 0$. The parameter λ is called a scale parameter. The parameter β is the shape parameter. The mean and variance of a Weibull random variable are the functions of the shape β and scale λ parameters can be expressed as $E(X) = \lambda \Gamma(1 + 1/\beta)$ and $Var(X) = \lambda^2 [\Gamma(1 + 2/\beta) - \Gamma^2(1 + 1/\beta)]$. We consider robust estimators which were proposed by [2] to achieve the robust estimates of the mean and variance of this distribution. The estimators proposed by [2] are robust to outliers, but they have the additional advantage of being an explicit function of the data.

In this study we restrict our attention to estimators that have the following set of properties: an explicit formula; a 50% breakdown point and a bounded IF. We propose the robust estimators of mean and variance by considering robust estimators based on Q, QLS and *Rmed*. We also derive their IFs and their breakdown points. The proposed estimators for mean and variance all have a high breakdown point and bounded IF.

In the following Section 3.1 and Section 3.2, quantile and regression estimators are given for robust mean and variance estimators of Weibull distribution.

3.1. Quantile-estimators. The quantile estimators of mean and variance for Weibull distribution are given by

$$\begin{aligned} \hat{\mu}_{\mathrm{WQ}} &= \hat{\lambda}_{\mathrm{Q}} \Gamma(1 + 1/\hat{\beta}_{\mathrm{Q}}), \\ \hat{\sigma}_{\mathrm{WQ}} &= \hat{\lambda}_{\mathrm{Q}}^2 [\Gamma(1 + 2/\hat{\beta}_{\mathrm{Q}}) - \Gamma^2(1 + 1/\hat{\beta}_{\mathrm{Q}})] \end{aligned} \qquad (3.1)$$

where [2] proposed the quantile estimators of shape and scale parameter
$\hat{\beta}_{\mathrm{Q}} = \frac{1}{\log(\hat{q}_{\alpha_2}/\hat{q}_{\alpha_1})} \log \frac{\log(1-\alpha_2)}{\log(1-\alpha_1)}$, and $\hat{\lambda}_{\mathrm{Q}} = \hat{q}_{\alpha} / [-\log(1-\alpha)]^{1/\hat{\beta}_{\mathrm{Q}}}$.

3.2. Regression estimators. The quantiles of the general log-Weibull distribution in $G^{-1}_{\lambda,\beta}(\alpha) = \beta^{-1}\log(-\log(1-\alpha))+\log\lambda$ are linearly related to the quantiles of the standard log-Weibull distribution, with intercept $b_0 = \log\lambda$ and slope $b_1 = 1/\beta$. Replacing the theoretical quantiles with their empirical counterparts yields a linear regression equation $y_i = b_0 + b_1 z_i + \varepsilon_i$ where $y_i = \log\hat{q}_{i/(n+1)}$ and $z_i = G^{-1}(i/(n+1))$. [2] considered two robust and explicit regression estimators for b_1 and b_0: the *Quantile Least Squares* and the *Repeated Median* estimators. The corresponding estimates of scale and shape of the Weibull distribution were then directly given by $\hat{\lambda} = \exp(\hat{b}_0)$ and $\hat{\beta} = 1/\hat{b}_1$.

Quantile Least Square: The *QLS* estimators of mean and variance of Weibull distribution are given by

$$\begin{aligned} \hat{\mu}_{\mathrm{WQLS}} &= \hat{\lambda}_{\mathrm{QLS}}\Gamma(1+1/\hat{\beta}_{\mathrm{QLS}}) \\ \hat{\sigma}_{\mathrm{WQLS}} &= \hat{\lambda}^2_{\mathrm{QLS}}[\Gamma(1+2/\hat{\beta}_{\mathrm{QLS}}) - \Gamma^2(1+1/\hat{\beta}_{\mathrm{QLS}})] \end{aligned} \tag{3.2}$$

where the *QLS* estimators of shape and scale parameters proposed by [2] :
$\hat{\lambda}_{\mathrm{QLS}} = \exp(\hat{b}_{0\mathrm{QLS}})$ and $\hat{\beta}_{\mathrm{QLS}} = 1/\hat{b}_{1\mathrm{QLS}}$ where $\hat{b}_{0\mathrm{QLS}}$ and $\hat{b}_{1\mathrm{QLS}}$ are *QLS* regression estimators for b_0 and b_1 (for details see [2]).

Repeated Median: The *Rmed* estimators of mean and variance of Weibull distribution are given by

$$\begin{aligned} \hat{\mu}_{\mathrm{WRmed}} &= \hat{\lambda}_{\mathrm{Rmed}}\Gamma(1+1/\hat{\beta}_{\mathrm{Rmed}}) \\ \hat{\sigma}_{\mathrm{WRmed}} &= \hat{\lambda}^2_{\mathrm{Rmed}}[\Gamma(1+2/\hat{\beta}_{\mathrm{Rmed}}) - \Gamma^2(1+1/\hat{\beta}_{\mathrm{Rmed}})] \end{aligned} \tag{3.3}$$

where the *Rmed* estimators of shape and scale parameters were proposed by [2]: $\hat{\lambda}_{\mathrm{Rmed}} = \exp(\hat{b}_{0\mathrm{Rmed}})$ and $\hat{\beta}_{\mathrm{Rmed}} = 1/\hat{b}_{1\mathrm{Rmed}}$ where $\hat{b}_{0\mathrm{Rmed}}$ and $\hat{b}_{1\mathrm{Rmed}}$ are *Rmed* regression estimators for b_0 and b_1 (for details see [2]).

4. Robustness of estimators

Robustness of estimators can be measured in different ways. The most important robustness measures are IF and BP of the estimators. In this study we derive IF and BP for the proposed estimators of mean and variance for Weibull distribution. The IF acts like the first derivative of functional defined on empirical distributions which we evaluate at the estimator. IF should be bounded to be robust. The breakdown point is a global robustness measure which describes how many percent gross errors are still tolerated before increasingly offensive outliers force our estimator to wander off to infinity. In next Section 4.1 and Section 4.2 we derive their IF and then their breakdown points, respectively.

4.1. IFs for Proposed Robust Estimators. IF gives price information about how to respond to a small amount of distortion at any point. Naturally, the estimators are very sensitive form of the distribution F, too much affected by deterioration in small quantities.

The statistical functionals corresponding with the mean and variance robust estimators are given by

$$\mu_{\mathrm{RE}}(F) = \lambda_{\mathrm{RE}}(F)\Gamma(1+1/\beta_{\mathrm{RE}}(F)) \tag{4.1}$$

$$\sigma_{\mathrm{RE}}(F) = \lambda^2_{\mathrm{RE}}(F)[\Gamma(1+2/\beta_{\mathrm{RE}}(F)) - \Gamma^2(1+1/\beta_{\mathrm{RE}}(F))]. \tag{4.2}$$

The IF of the functional $\mu_{\rm RE}$ at the Weibull distribution $F_{\lambda,\beta}$ in (4.1) is given by

$$\begin{aligned} IF(x_0;\mu_{\rm RE},F_{\lambda,\beta}) &= \frac{\partial}{\partial\varepsilon}(\mu_{\rm RE}(F_\varepsilon))|_{\varepsilon=0} \\ &= \Gamma(1+1/\hat{\beta}_{\rm RE})\Big(IF(x_0;\lambda_{\rm RE},F_{\lambda,\beta}) \\ &- \frac{\hat{\lambda}_{\rm RE}}{\hat{\beta}^2_{\rm RE}}\psi(1+1/\hat{\beta}_{\rm RE})IF(x_0;\beta_{\rm RE},F_{\lambda,\beta})\Big). \end{aligned} \tag{4.3}$$

The IF of the functional $\sigma_{\rm RE}$ at the Weibull distribution $F_{\lambda,\beta}$ in (4.2) is given by

$$\begin{aligned} IF(x_0;\sigma_{\rm RE},F_{\lambda,\beta}) &= \frac{\partial}{\partial\varepsilon}(\sigma_{\rm RE}(F_\varepsilon))|_{\varepsilon=0} \\ &= 2\hat{\lambda}_{\rm RE}IF(x_0;\lambda_{\rm RE},F_{\lambda,\beta})[\Gamma(1+2/\hat{\beta}_{\rm RE})-\Gamma(1+1/\hat{\beta}_{\rm RE})^2] \\ &+ 2\frac{\hat{\lambda}^2_{\rm RE}}{\hat{\beta}^2_{\rm RE}}IF(x_0;\beta_{\rm RE},F_{\lambda,\beta})\Big[-\psi(1+2/\hat{\beta}_{\rm RE})\Gamma(1+2/\hat{\beta}_{\rm RE}) \\ &+ \Gamma(1+1/\hat{\beta}_{\rm RE})^2\psi(1+1/\hat{\beta}_{\rm RE})\Big]. \end{aligned} \tag{4.4}$$

where $\hat{\beta}_{\rm RE}$ and $\hat{\lambda}_{\rm RE}$ are the shape and scale robust Q, QLS and $Rmed$ estimators of Weibull parameters. For IFs $IF(x_0;\lambda_{\rm RE})$ and $IF(x_0;\beta_{\rm RE})$ in (4.3) and (4.4), see [2].

The IFs for the classic and robust estimators of mean and variance are pictured in Figure 1. It is seen that while the IF of least square (LS) estimator is unbounded function, the IFs for robust estimators are bounded. It should be considered that the IFs of quantile mean and variance estimator are step functions. As a result we can say that the proposed estimators are B-robust which means that their IFs are bounded.

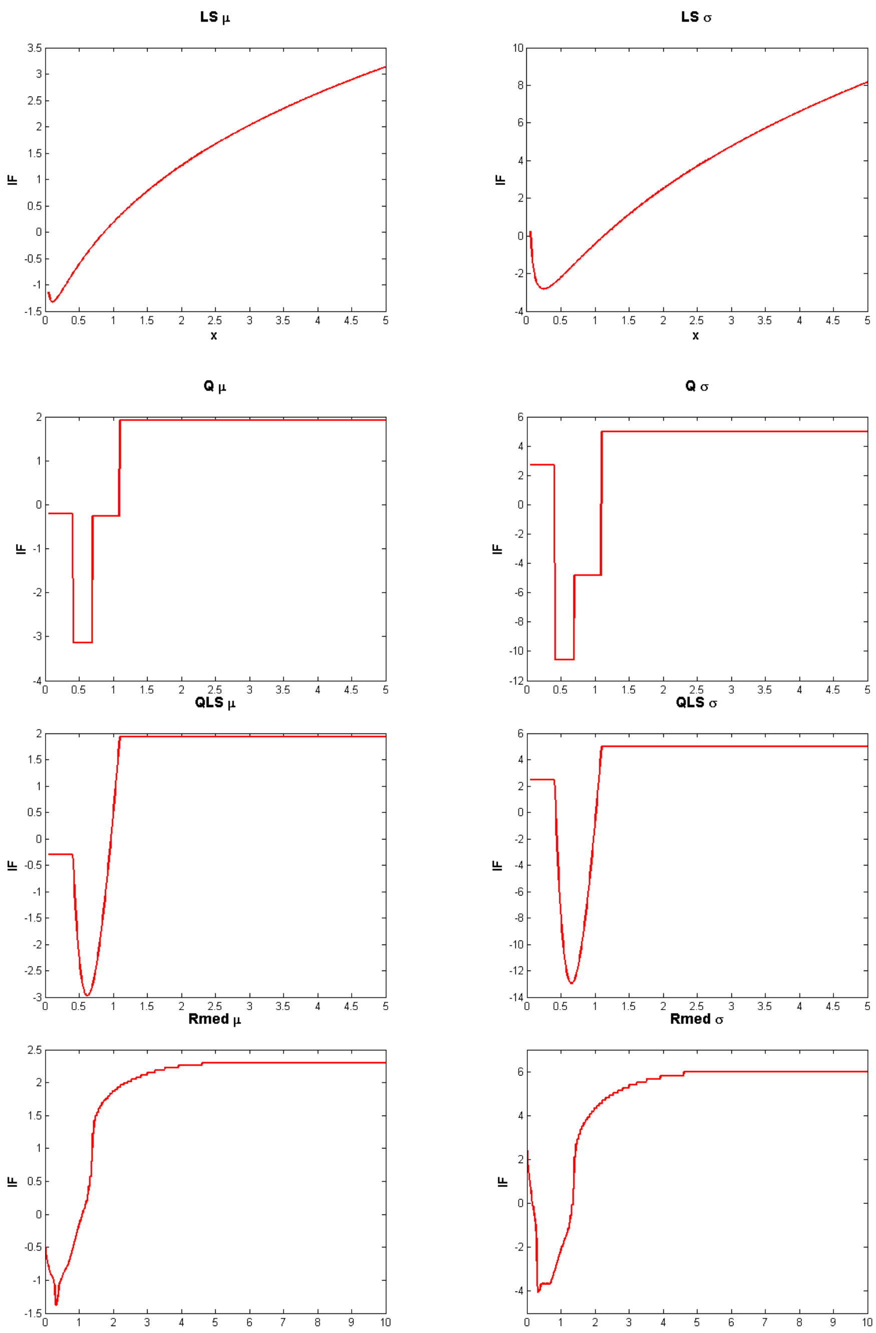
4.2. Breakdown Points of Proposed Robust Estimators. The breakdown point of an estimator is the proportion of incorrect observations an estimator can handle before given an arbitrarily large result. The higher the breakdown point of an estimator, the more robust it is. Instinctively, we can understand that a breakdown point can not exceed %50 because if more than half of the observations are contaminated, it is not possible to distinguish between the underlying distribution and the contaminating distribution. Therefore, the maximum breakdown point is 0.5 and there are estimators which achieve such a breakdown point.

The breakdown points of robust estimators were examined earlier in the previous studies examined by some authors: For linear regression parameters least square estimators: α and repeated median estimators: %50. For shape and scale estimators based on Q, QLS, $Rmed$ methods : [2]. To characterize the robustness of the proposed estimators, we derive their BP, defined as the smallest proportion of observations (for $n\to\infty$) that needs to be replaced with arbitrary values in order for the estimation of λ or β to be arbitrarily close to zero (implosion) or infinity (explosion). To define the breakdown point of the mean and variance we consider the BP of shape and scale estimators of Weibull distribution.

The BP of the mean estimator for Weibull distribution is given by

$$\varepsilon^*_n(\mu,F_n) = min\{\varepsilon^+_n(\mu,F_n),\varepsilon^-_n(\mu,F_n)\}, \tag{4.5}$$

Figure 1. IF of the mean and variance estimators of Weibull Distribution



In 4.5 the explosion BP is

$$\varepsilon_n^+(\mu, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \sup_{F'_n} M(\mu(F'_n)) = \infty\}. \tag{4.6}$$

We get $\sup_{F'_n} M(\mu(F'_n)) = \infty$, if $\lambda \to \infty$ or $\gamma(1+1/\beta) \to \infty$. For $\lambda \to \infty$ the BP is $\varepsilon^+(\lambda, F)$. For $\gamma(1+1/\beta) \to \infty$, if $\beta = -1, -1/2, -1/3, -1/4$. In this condition there is no BP since β is not going to infinity or zero. Therefore the explosion BP is $\varepsilon^+(\mu, F) = (\varepsilon^+(\lambda, F))$.

In 4.5 the implosion BP is

$$\varepsilon_n^-(\mu, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \inf_{F'_n} M(\mu(F'_n)) = 0\}. \tag{4.7}$$

We get $\inf_{F'_n} M(\mu(F'_n)) = 0$, if $\lambda \to 0$ or $\gamma(1+1/\beta) \to 0, \varepsilon^+(\beta, F)$. For $\lambda \to 0$ the BP is $\varepsilon^-(\lambda, F)$. For $\gamma(1+1/\beta) \to 0, \varepsilon^+(\beta, F)$: if $\beta \to \infty$, $\gamma(1+1/\beta) = 1/\beta\gamma(1/\beta) \to 0$. So for $\beta \to \infty$ the BP is $\varepsilon^+(\beta, F)$. The implosion BP is $\varepsilon^-(\mu, F) = (\varepsilon^-(\lambda, F), \varepsilon^+(\beta, F))$. As a result the BP of the mean estimator is given by

$$\begin{aligned} \varepsilon_n^*(\mu, F_n) &= \min\{\varepsilon_n^+(\mu, F_n), \varepsilon_n^-(\mu, F_n)\} \\ &= \min\{\varepsilon^+(\lambda, F), \varepsilon^-(\lambda, F), \varepsilon^+(\beta, F)\} \end{aligned}$$

The BP of variance estimator of Weibull distribution is given by

$$\varepsilon_n^*(\sigma, F_n) = min\{\varepsilon_n^+(\sigma, F_n), \varepsilon_n^-(\sigma, F_n)\}. \tag{4.8}$$

In 4.8 the explosion BP is $\varepsilon_n^+(\sigma, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \sup_{F'_n} M(\sigma(F'_n)) = \infty\}$ We get $\sup_{F'_n} M(\sigma(F'_n)) = \infty$,if $\lambda \to \infty$ or $\gamma(1+2/\hat{\beta}) - \gamma(1+1/\hat{\beta}) > 0$. For $\lambda \to \infty$ can be obtained if $(\varepsilon_n^+(\lambda, F))$. For $\gamma(1+2/\hat{\beta}) - \gamma(1+1/\hat{\beta}) > 0$ can be obtained if $\beta > 0$. So the BP is $\varepsilon_n^+(\beta, F)$. Therefore the explosion BP of variance estimator is obtained $\varepsilon^+(\sigma, F) = (\varepsilon^+(\lambda, F), \varepsilon^+(\beta, F))$.

In 4.8 the implosion BP is $\varepsilon_n^+(\sigma, F_n) = min\{\frac{m}{n}, m \in 1, \ldots, n | \inf_{F'_n} M(\sigma(F'_n)) = 0\}$ We get $\inf_{F'_n} M(\sigma(F'_n)) = 0$, if $\lambda \to 0$ or $\beta \to \infty$. For $\lambda \to 0$, the BP is $\varepsilon^-(\lambda, F)$. For $\beta \to \infty$ the BP is$\varepsilon^+(\beta, F)$. Therefore the implosion BP of variance estimator is obtained $\varepsilon^-(\sigma, F) = (\varepsilon^-(\lambda, F), \varepsilon^+(\beta, F))$.

As a result the BP of variance estimator is given by

$$\begin{aligned} \varepsilon_n^*(\sigma, F_n) &= \min\{\varepsilon_n^+(\sigma, F_n), \varepsilon_n^-(\sigma, F_n)\} \\ &= \min\{\varepsilon^+(\lambda, F), \varepsilon^+(\beta, F), \varepsilon^-(\lambda, F)\} \end{aligned}$$

Asymptotic BPs for the robust estimators of mean and variance of Weibull distribution are given by Table 1.

Table 1. Asymptotic BP for the robust estimators for mean and variance of Weibull distribution

Method	$\varepsilon^*(\mu)$	$\varepsilon^*(\sigma)$
Q	$\min(\alpha, 1-\alpha, \alpha_1, \alpha_2-\alpha_1)$ $\alpha \neq 1-e^{-1}$ $\min(\alpha, 1-\alpha)$ $\alpha = 1-e^{-1}$	$\min(\alpha, 1-\alpha, \alpha_1, \alpha_2-\alpha_1)$ $\alpha \neq 1-e^{-1}$ $\min(\alpha, 1-\alpha, \alpha_2-\alpha_1)$ $\alpha = 1-e^{-1}$
QLS	$\min(\overline{\alpha}, 1-2\overline{\alpha})$	$\min(\overline{\alpha}, 1-2\overline{\alpha})$
$Rmed$	%50	%50

5. Simulation Study

The behavior of the robust Welch test statistic is examined according to all three methods with 10,000 repetitions. The type I errors of proposed test statistic is obtained according to robust methods by considering various experimental designs. At the end of the simulation study robust test statistic will be compared in terms of the type I errors, and the comments will be made for experimental designs.

Since the mean and variance of Weibull distribution are functions of the shape and scale parameters, the creation of different combinations depends only on the parameters of the distribution. When scale parameter is one and shape parameter takes different values, the mean and variance do not change much. However, when the scale parameter value is changed, the mean and variance change a lot. Therefore, to generate different experimental design a scale parameter is fixed, it takes $\lambda = 1$ with different values of shape parameter. For example when the shape parameter β is equal to one, then this distribution reduces to the exponential distribution. A model that results in values of probability $prob\{y \geq E(y)\}$ substantially greater or smaller than 0.5 is hardly of any practical interest. For the values of β less than 1.2, $prob\{y \geq E(y)\} < 0.4$ [11]. Moreover, [4] argue that in most applications where a Weibull distribution is applicable β is greater than one. For these reasons, we consider values of $\beta \geq 1.5$. In the simulation study, the value of the shape parameters are selected as in Table 2 with respect to the different experimental designs which we want to create. In this table for equal means, homogeneous variances A1, B1 are used for balanced sample size, C1, D1 are used for unbalanced sample size. For unequal means, heterogenous variances A2, B2 are used for balanced sample size and C2, D2 are used for unbalanced sample size. As we mentioned before to generate data with equal means + homogeneous variances and unequal means + heterogenous variances, we just change the shape parameter of Weibull distribution.

Table 2. Experimental designs for k=3 and k=6

		$k=3$			$k=6$					
A1	n_i	5	5	5	5	5	5	5	5	5
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
A2	n_i	5	5	5	5	5	5	5	5	5
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5
B1	n_i	10	10	10	10	10	10	10	10	10
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
B2	n_i	10	10	10	10	10	10	10	10	10
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5
C1	n_i	5	10	15	5	10	15	5	10	15
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
C2	n_i	5	10	15	5	10	15	5	10	15
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5
D1	n_i	10	20	30	10	20	30	10	20	30
	β	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5	1.5
D2	n_i	10	20	30	10	20	30	10	20	30
	β	1.5	2	2.5	1.5	2	2.5	1.5	2	2.5

In the simulation study the reference distribution is $W(1,\beta)$ whose characteristic is mentioned in Table 2. By using the proposed estimators Q, QLS and $Rmed$, the type I errors of the robust test statistic and classical ANOVA are obtained with 10,000 repetitions at the significance level of 5%. For quantile methods, $\alpha = 30\%$ is taken. Four different models are discussed below to test the behaviors of the test statistic, when the model is deteriorated and in the presence of outliers:

- Model 1: Clean sample (Reference distribution $W(1,\beta)$),
- Model 2: Dixon model (n-1 observations from $W(1,\beta)$, 1 observation from $W(2,\beta)$),
- Model 3: Mixture model ($0.80W(1,\beta)$ + 0.20 $W(2,\beta/2)$),
- Model 4: Contaminated model (0.80 $W(1,\beta)$ + 0.20(100 Uniform(0,1))).

We obtain $\tau = \frac{\sum_{i=1}^{M} F_{H_i} > F_{T_i}}{M} * 100$ value with respect the classical ANOVA and RW test statistics with 10,000 repetitions. In this equation F_{H_i} indicates the calculated test statistic and F_{T_i} indicates the F table value at the significance of 5% for the ith simulation, so desirable value of τ is to be close $\tau \cong 5$.
By considering combinations of the above-mentioned trial simulation study, the type I errors of test statistic based on three methods is obtained and the results (type I error * 100=$\boldsymbol{\tau}$) are given. The robust test statistic will be compared in terms of type I errors and the comments in detail for each trial will be made.

The results of $\boldsymbol{\tau}$ values for experimental designs with equal means and homogeneous variances are given in Table 3 for $k = 3$. While the classical ANOVA does not deteriorate for clean model (model 1), it badly deteriorates for contaminated model especially for unbalanced sample size. As seen from this table, the results of Q methods are not good. The type I errors of RW test statistic based on $Rmed$ methods are desired level especially for experimental design C1 and D1. RW test statistic based on QLS method can be an alternative only for experimental design C1.

Table 3. The $\boldsymbol{\tau}$ values for k=3, Equal means, homogeneous variances

RE	Model	A1		B1		C1		D1	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	4.80	6.71	5.16	11.8	6.36	11.03	6.37	14.37
	2	3.90	10.55	4.27	13.22	5.53	15.99	6.17	17.68
	3	3.10	9.27	3.61	10.80	4.54	12.98	5.40	15.55
	4	1.60	9.70	0.60	13.90	28.85	11.19	83.61	15.75
QLS	1	5.14	6.91	5.08	14.45	6.23	7.93	6.79	18.42
	2	3.75	2.45	4.23	5.38	5.32	7.89	6.21	8.20
	3	3.11	3.85	3.61	10.40	4.44	5.45	5.58	15.89
	4	1.60	9.70	0.08	18.99	29.85	5.87	82.85	14.91
Rmed	1	4.53	8.96	4.86	6.07	6.64	13.91	6.19	7.00
	2	3.89	7.08	4.46	5.66	5.02	6.86	5.90	7.72
	3	3.26	5.90	3.63	4.43	4.76	6.22	5.37	4.93
	4	0.14	4.51	0.04	4.93	30.25	3.79	82.50	4.88

The results of $\boldsymbol{\tau}$ values for experimental designs with non-equal means and heterogenous variance are given in Table 4 for $k = 3$. For unbalanced sample size classical ANOVA is deteriorate for all methods. Only the RW test statistic based on $Rmed$ has good performance for contaminated model. But RW test statistic does not work for other robust

Table 4. The $\boldsymbol{\tau}$ values for k=3, unequal means, heterogenous variances

RE	Model	A2		B2		C2		D2	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	5.30	7.21	6.35	10.92	12.32	12.46	12.39	14.62
	2	4.74	12.35	4.95	14.32	11.35	18.35	12.05	18.65
	3	3.73	11.35	4.07	12.91	9.21	16.71	9.83	16.35
	4	0.16	9.70	0.06	13.56	32.66	11.37	83.47	14.96
QLS	1	5.52	8.44	5.20	15.99	12.13	16.71	12.55	18.78
	2	4.60	3.76	4.85	5.81	9.04	13.21	12.05	8.93
	3	3.90	5.19	4.05	12.87	8.75	14.08	10.22	17.25
	4	0.17	4.89	0.05	12.91	31.96	12.41	83.74	14.43
Rmed	1	5.89	8.54	5.47	6.61	12.20	9.33	12.63	7.19
	2	4.74	8.23	5.38	5.95	11.25	10.17	12.08	7.98
	3	4.12	6.22	4.41	4.68	8.78	7.77	9.31	8.00
	4	0.15	4.04	0.04	4.29	32.09	4.20	83.37	5.11

methods. For clean, and very few corrupted samples the Type I error level of classical ANOVA is considerably good since the variances are homogeneous for D1 experimental design. However for contaminated model classical ANOVA is deteriorated badly.

The results of $\boldsymbol{\tau}$ values for experimental designs with equal means and homogeneous variances are given in Table 5 for $k = 6$. As seen from the results, the *RW* test statistic based on *Rmed* robust method works well for mixture and contaminated model in only A2 and B2 experimental designs, but the other robust methods do not work.

Table 5. The $\boldsymbol{\tau}$ values for k=6, Equal means, homogeneous variances

RE	Model	A1		B1		C1		D1	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	4.70	10.61	4.92	16.91	5.89	17.73	6.13	23.98
	2	3.90	10.55	4.03	24.89	4.72	27.97	5.43	23.12
	3	3.20	20.47	3.74	22.02	4.07	24.93	4.60	26.45
	4	0.9	22.12	0.77	24.45	61.31	24.07	86.55	28.09
QLS	1	6.03	17.21	4.74	29.80	13.00	35.22	5.65	33.07
	2	3.97	5.15	4.25	9.31	4.53	34.30	5.14	13.73
	3	3.28	9.22	3.28	24.20	4.13	26.12	4.17	28.91
	4	0.13	11.47	0.95	14.19	60.99	28.44	85.85	30.00
Rmed	1	5.10	25.22	5.10	9.50	5.67	27.70	6.18	10.71
	2	3.65	15.03	4.34	9.69	4.96	12.59	5.07	10.65
	3	3.10	11.98	3.39	6.73	4.55	21.27	4.34	7.20
	4	0.14	10.33	0.94	7.10	61.50	7.81	85.69	7.70

The results of $\boldsymbol{\tau}$ values for experimental designs with non-equal means and heterogenous variance are given in Table 6 for $k = 6$. As you see, we can conclude the results for contaminated model for $(k = 6)$ such as that the *RW* test statistic only based on *Rmed* robust method gives desirable results.

To sum up all results, we can say that in the case of contamination proposed robust Welch test statistic can be used for ($k = 3$). When the number of group is small, for contaminated models the Type I errors of RW test statistic has good performance. The number of group is small *Rmed* according to the methods of RW test statistic Type I errors is desirable. The number of group grows, RW test statistic has undergone distortion.

Table 6. The $\boldsymbol{\tau}$ values for k=6, unequal means, heterogenous variances

RE	Model	A2		B2		C2		D1	
		F	**RW**	**F**	**RW**	**F**	**RW**	**F**	**RW**
Q	1	5.94	11.72	5.50	30.62	12.73	19.55	13.12	22.71
	2	4.52	24.98	5.24	29.98	11.25	31.71	12.01	29.81
	3	3.41	22.49	4.09	22.43	9.09	27.71	9.48	26.61
	4	0.90	22.12	0.81	24.45	62.43	24.71	86.52	27.93
QLS	1	4.67	20.41	6.12	31.18	13.00	30.06	13.04	31.90
	2	4.53	6.81	5.07	9.81	11.31	13.87	12.10	12.71
	3	3.37	12.73	3.60	26.59	8.47	26.54	9.77	29.36
	4	0.14	10.97	1.13	27.83	62.78	23.09	86.80	26.67
Rmed	1	5.24	16.41	6.07	10.22	12.61	13.52	12.96	10.31
	2	4.38	16.28	4.93	9.65	10.73	15.45	11.90	11.38
	3	3.62	13.25	3.55	7.71	8.97	11.64	9.27	10.28
	4	0.18	9.46	0.91	6.37	62.51	8.78	86.41	6.69

6. Conclusion

The purpose of this study is to develop test statistic for one-way ANOVA by using robust methods under Weibull distribution with outlier. For this purpose, we propose the robust estimators for mean and variance of Weibull distribution. We also derive not only their BP but also their IFs. The proposed estimators for mean and variance all have a high BP and bounded IF. RW test statistic is obtained by using the estimators based on Q, QLS and *Rmed*. The behavior of the modified robust test statistic is examined by simulation study.

In the simulation study, using various experimental designs, type I errors of the improved robust test statistic and classical ANOVA under Weibull distribution are obtained with respect to three different robust estimators. Balanced and unbalanced sample sizes for k=3,6 groups with homogeneous and heterogeneous variances are considered. Then the simulation results show up: For unbalanced sample size classical ANOVA is deteriorated. When the number of groups is small ($k = 3$), the RW test statistic based on *Rmed* and QLS methods performance is not deteriorated badly. When the number of groups is increasing, especially for contaminated models the proposed RW test based on *Rmed* method gives desirable results. The RW test statistic based on *Rmed* has the best performance for all experimental design especially in contaminated models, while the RW test statistic based on Q does not work. QLS method can be used as an alternative to *Rmed* method.

References

[1] Abernethy, R.B., Breneman, J.E., Medlin, C.H. and Reinman, G.L., 1983, *Weibull Analysis Handbook*, Technical Report, AFWAL-TR-83-207, Air Force Wright Aeronautical Laboratories, Washington, D.C., http://handle.dtic.mil/100.2/ADA143100.

[2] Boudt, K., Caliskan, D. and Croux, C., 2011, *Robust explicit estimators of Weibull parameters*, Metrika, 73, 187-209.

[3] Cochran, W.G., 1937, *Problems arising in the analysis of a series of similar experimentals*, Journal of Royal Statistics Soc.Supp., 4, 102-118.

[4] Cohen, A. C. and Whitten, B. J., 1988, *Parameter Estimation in Reliability and Life Span Models*, New York: Marcel Dekker.

[5] DerSimonian, R. and Laird, N., 1986, *Meta-analysis in clinical trials*, Controlled Clin. Trials, 7, 177-188.

[6] Gupta, R.D. and Kundu, D., 2001, *Exponentiated Exponential Family: An Alternative to Gamma and Weibull Distributions*, Biometrical Journal 43,1, 117-130

[7] James, G., S., 1951, *The comparison of several groups of observations when the ratios of the population variances are unkown*, Biometrika, 38, 324-329.

[8] Karagöz, D., 2012, *Robust Analaysis of Variance*, PhD thesis: Hacettepe University, Ankara, Turkey.

[9] Kulinskaya, E., Staudte, R.G. and Gao, H., 2003, *Power approximations in testing for unequal means in a One-Way ANOVA weighted for unequal variances*, Communications in Statistics Theory and Methods, 32,2353-2371.

[10] Kulinskaya, E. and Dollinger, M.B., 2007, *Robust weighted One-Way ANOVA: Improved approximation and efficiency*, Journal of Statistical Planning and Inference, 137, 462-472.

[11] Şenoğlu, B., 2005, *Robust 2^k Factorial Design with Weibull Error Distributions*, Journal of Applied Statistics, 32-10, 1051-1066.

[12] Takkouche, B., Cadarso-Suarez, C. and Spiegelman, D., 1999, *Evaluation of old and new tests of heterogeneity in epidemiologic meta-analysis*, Amer.J. Epidemiology 150, 206-215.

[13] Welch, B.L., 1951, *On the comparison of several mean values: an alternative approach.* Biometrika, 38, 330-336.

[14] Wilcox, R.R., 1995, *The practical importance of heteroscedastic methods, using trimmed means versus means, and designing simulation studies*, British J. Math.Statist.Psych., 48, 99-114.

[15] Wilcox, R.R., 1997, *Introduction to Robust Estimation and Hypothesis Testing*, Academic Press, New York.