ROBUST ESTIMATION AND HYPOTHESIS TESTING IN 2^k FACTORIAL DESIGN

BIRDAL ŞENOĞLU

ABSTRACT. The topic of this work is an extension of our previous work on robust 2^k factorial design with Weibull error distributions. In this paper, we obtain robust and efficient estimators of the parameters in the 2^k factorial design by using the methodology known as modified maximum likelihood (MML) and propose new test statistics based on MML estimators for testing the main effects and the interactions when the distribution of error terms is generalized logistic. We show that the proposed test statistics are more powerful and robust than the traditional test statistics based on the least squares (LS) estimators.

1. INTRODUCTION

Factorial designs are the most efficient designs in terms of time and cost when we evaluate two or more factors simultaneously. They provide us information about the main effects and the interactions among the various factors, Fisher [5] and Yates [14]. 2^k factorial design is the simplest type of factorial designs. Here, k represents the number of factors and 2 represents the number of levels for each factor. These levels are usually referred to as "low" and "high" levels. 2^k factorial designs are very useful for preliminary exploration when there are large numbers of factors in a factorial design. They are used very widely in agricultural experimentations, in engineering experimentations, etc.

Şenoğlu [10] considered the 2^k factorial design when the distribution of error terms is Weibull $W(p, \sigma)$. From the methodology of modified likelihood, they developed robust and efficient estimators for the parameters in 2^k factorial design. F statistics based on MML estimators for testing main effects and interactions were defined. They were shown to have high powers and better robustness properties as compared to the normal theory solutions.

In this study, we extend the results to the case where the distribution of error terms are independent and identically distributed (iid) according to a generalized logistic distribution. The family of generalized logistic distribution $GL(b, \sigma)$ is given by

$$\frac{b}{\sigma} \frac{\exp(-e/\sigma)}{\{1 + \exp(-e/\sigma)\}^{b+1}}, \quad -\infty < e < \infty.$$

The cumulative distribution is given by

$$F(e) = (1 + \exp(-e/\sigma))^{-b}. \tag{1.2}$$

One of the main motivation of this work is that this family beautifully complements the family of Weibull distributions, (i) its support is on IR: $(-\infty, \infty)$ and (ii) its represents leptokurtic distributions ($\beta_2 > 3$) while most Weibull distributions are mesokurtic ($\beta_2 < 3$). See the following table for the skewness ($\sqrt{\beta_1}$) and kurtosis (β_2) values of the $GL(b, \sigma)$ distribution.

$b =$	0.5	**1**	2	4	6
$\sqrt{\beta_1}$	−0.86	0	0.33	0.75	0.92
β_2	5.40	4.20	4.33	4.76	4.95

It may be noted that for $b=1$, $GL(b, \sigma)$ reduces to the well-known logistic distribution which has been, in many studies, used as a substitute for a normal distribution; see, for example, Berkson [3].

2. THE 2^3 DESIGN

It should be noted that 2^2 design is a special case of Şenoğlu and Tiku [11], because two-way classification model reduces to 2^2 factorial design when $i = j = 1, 2$. Therefore, we will not pursue it in this study for the sake of brevity. Let's consider the case where there are three factors (say A, B and C), each of which has two levels, i. e., 2^3 factorial design. The model for such an experiment is

$$y_{ijkl} = \mu + \tau_i + \beta_j + \gamma_k + (\tau\beta)_{ij} + (\tau\gamma)_{ik} + (\beta\gamma)_{jk} + (\tau\beta\gamma)_{ijk} + e_{ijkl} \tag{2.1}$$

(i=1, 2; j=1, 2; k =1, 2; l=1, 2,….., n)

where $-\infty < \mu < \infty$ is the general or overall mean common to all the observations. τ_i, β_j and γ_k are the effects due to the ith level of factor A, jth level of factor B and kth level of factor C, respectively. $(\tau\beta)_{ij}$, $(\tau\gamma)_{ik}$ and $(\beta\gamma)_{jk}$ are the effects of the two-factor interactions between τ_i and β_j, τ_i and γ_k and β_j and γ_k, respectively.

$(\tau\beta\gamma)_{ijk}$ is the effect of three-factor interaction between τ_i, β_j and γ_k and e_{ijkl} is the random error associated with the lth observation, at the ith level of the factor A, jth level of factor B and kth level of factor C. The factors A, B and C are considered as fixed and the designs are assumed to be completely randomized in the rest of the paper.

3. THE MML ESTIMATORS

Let $z_{ijk(l)} = (y_{ijk(l)} - \mu - \tau_i - \beta_j - \gamma_k - (\tau\beta)_{ij} - (\tau\gamma)_{ik} - (\beta\gamma)_{jk} - (\tau\beta\gamma)_{ijk})/\sigma$ $(1 \le i \le 2,\ 1 \le j \le 2,\ 1 \le k \le 2,\ 1 \le l \le n)$ be the ordered variates, where $y_{ijk(1)} \le y_{ijk(2)} \le \ .... \le y_{ijk(n)}$ are the ordered statistics obtined by arranging the random observations in the ith level of the factor A, jth level of factor B and kth level of factor C, i. e. y_{ijkl}, in ascending order of magnitude. The likelihood function and the log-likelihood function are

$$L \propto \left(\frac{1}{\sigma}\right)^N \frac{e^{-\sum_i\sum_j\sum_k\sum_l z_{ijk(l)}}}{\prod_i\prod_j\prod_k\prod_l [1 + e^{-z_{ijk(l)}}]^{b+1}} \tag{3.1}$$

and

$$\ln L \propto -N \ln\sigma - \sum_i\sum_j\sum_k\sum_l z_{ijk(l)} - (b+1)\sum_i\sum_j\sum_k\sum_l \ln\left[1 + e^{-z_{ijk(l)}}\right], \tag{3.2}$$

respectively. Since complete sums are invariant to ordering, i. e., $\sum_{i=1}^{n} f(y_i) = \sum_{i=1}^{n} f(y_{(i)})$ where $f(y)$ is any function of y. By using equation (3.2), we obtain the following likelihood equations for each model parameters in (2.1)

$$\frac{\partial \ln L}{\partial \mu} = \frac{N}{\sigma} - \frac{(b+1)}{\sigma}\sum_i\sum_j\sum_k\sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}},$$

$$\frac{\partial \ln L}{\partial \tau_i} = \frac{bcn}{\sigma} - \frac{(b+1)}{\sigma}\sum_j\sum_k\sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}},$$

$$\frac{\partial \ln L}{\partial \beta_j} = \frac{acn}{\sigma} - \frac{(b+1)}{\sigma}\sum_i\sum_k\sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}},$$

$$\frac{\partial \ln L}{\partial \gamma_k} = \frac{abn}{\sigma} - \frac{(b+1)}{\sigma}\sum_i\sum_j\sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}},$$

$$\frac{\partial \ln L}{\partial (\tau\beta)_{ij}} = \frac{cn}{\sigma} - \frac{(b+1)}{\sigma}\sum_k\sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}}, \tag{3.3}$$

$$\frac{\partial \ln L}{\partial (\tau\gamma)_{ik}} = \frac{bn}{\sigma} - \frac{(b+1)}{\sigma} \sum_j \sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}},$$

$$\frac{\partial \ln L}{\partial (\beta\gamma)_{jk}} = \frac{an}{\sigma} - \frac{(b+1)}{\sigma} \sum_i \sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}},$$

$$\frac{\partial \ln L}{\partial (\tau\beta\gamma)_{ijk}} = \frac{n}{\sigma} - \frac{(b+1)}{\sigma} \sum_l \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}}$$

and

$$\frac{\partial \ln L}{\partial \sigma} = -\frac{N}{\sigma} + \frac{1}{\sigma} \sum_i \sum_j \sum_k \sum_l z_{ijk(l)} - \frac{(b+1)}{\sigma} \sum_i \sum_j \sum_k \sum_l z_{ijk(l)} \frac{e^{-z_{ijk(l)}}}{1 + e^{-z_{ijk(l)}}}.$$

The likelihood equations in (3.3) do not yield explicit estimators of the model parameters because of the awkward function $g(z) = \frac{e^{-z}}{1+e^{-z}}$ and hence they must be solved by numerical methods. However, solving them by iterations is indeed problematic for reasons of (i) multiple roots, (ii) non-convergence of iterations, and (iii) convergence to wrong values; see, for example, Smith [9], Puthenpura and Sinha [8] and Vaughan [13]. Therefore, we linearize the term $g(z) = \frac{e^{-z}}{1+e^{-z}}$ by expanding it in a Taylor series around $t_{ijk(l)} = E(z_{ijk(l)})$, since $g(z)$ is almost linear in small intervals around $t_{ijk(l)}$. This methodology is known as modified maximum likelihood and was initiated by Tiku [12]. We then get

$$g(z_{ijk(l)}) \cong \alpha_l - \beta_l z_{ijk(l)} \tag{3.4}$$

where

$$\beta_l = \frac{e^{-t_{ijk(l)}}}{(1 + e^{-t_{ijk(l)}})^2} \quad \text{and} \tag{3.5}$$

$$\alpha_l = \frac{e^{-t_{ijk(l)}} + e^{-2t_{ijk(l)}} + e^{-t_{ijk(l)}} t_{ijk(l)}}{(1 + e^{-t_{ijk(l)}})^2} \quad (l = 1, 2, \ldots, n).$$

Exact values of $t_{ijk(l)}$ are available for $n \leq 15$ (see [1]) but, for convenience, we use their approximate values obtained from the equations

$$\int_{-\infty}^{t_{ijk(l)}} \frac{be^{-z}}{(1 + e^{-z})^{b+1}} dz = \frac{l}{n+1}, t_{ijk(l)} = -\ln\left(\left(\frac{l}{n+1}\right)^{-1/b} - 1\right). \tag{3.6}$$

Using approximate values instead of exact values does not adversely affect the efficiency of the MML estimators. Incorporating (3.4) in (3.3), the modified likelihood equations are obtained. The solutions of these equations are the following MML estimators;

$$\hat{\mu} = \hat{\mu}_{....} - (\Delta/m)\hat{\sigma}, \hat{\tau}_i = \hat{\mu}_{i...} - \hat{\mu}_{....}, \hat{\beta}_j = \hat{\mu}_{.j..} - \hat{\mu}_{....}, \hat{\gamma}_k = \hat{\mu}_{..k.} - \hat{\mu}_{....},$$

$$(\widehat{\tau\beta})_{ij} = \hat{\mu}_{ij..} - \hat{\mu}_{i...} - \hat{\mu}_{.j..} + \hat{\mu}_{....}, (\widehat{\tau\gamma})_{ik} = \hat{\mu}_{i.k.} - \hat{\mu}_{i...} - \hat{\mu}_{..k.} + \hat{\mu}_{....},$$

$$(\widehat{\beta\gamma})_{jk} = \hat{\mu}_{.jk.} - \hat{\mu}_{.j..} - \hat{\mu}_{..k.} + \hat{\mu}_{....}, \tag{3.7}$$

$$(\widehat{\tau\beta\gamma})_{ijk} = \hat{\mu}_{ijk.} - \hat{\mu}_{ij..} - \hat{\mu}_{i.k.} - \hat{\mu}_{.jk.} + \hat{\mu}_{i...} + \hat{\mu}_{.j..} + \hat{\mu}_{..k.} - \hat{\mu}_{....}$$

and

$$\hat{\sigma} = \frac{B + \sqrt{B^2 + 4NC}}{2\sqrt{N(N - 2^3)}}$$

where

$$m = \sum_{l=1}^{n} \beta_l, \Delta_l = \alpha_l - \frac{1}{b+1}; \quad \Delta = \sum_{l=1}^{n} \Delta_l, \quad \hat{\mu}_{....} = (1/8m)\sum_i \sum_j \sum_k \sum_l \beta_l y_{ijk(l)},$$

$$\hat{\mu}_{i...} = (1/4m)\sum_j \sum_k \sum_l \beta_l y_{ijk(l)}, \quad \hat{\mu}_{.j..} = (1/4m)\sum_i \sum_k \sum_l \beta_l y_{ijk(l)},$$

$$\hat{\mu}_{..k.} = (1/4m)\sum_i \sum_j \sum_l \beta_l y_{ijk(l)}, \quad \hat{\mu}_{ij..} = (1/2m)\sum_k \sum_l \beta_l y_{ijk(l)},$$

$$\hat{\mu}_{i.k.} = (1/2m)\sum_j \sum_l \beta_l y_{ijk(l)}, \quad \hat{\mu}_{.jk.} = (1/2m)\sum_i \sum_l \beta_l y_{ijk(l)},$$

$$\hat{\mu}_{ijk.} = (1/m)\sum_l \beta_l y_{ijk(l)}, \quad B = (b+1)\sum_i \sum_j \sum_k \sum_l (y_{ijk(l)} - \hat{\mu}_{ijk.})\Delta_l$$

and

$$C = (b+1)\sum_i \sum_j \sum_k \sum_l (y_{ijk(l)} - \hat{\mu}_{ijk.})^2 \beta_l.$$

The divisor N in the expression for $\hat{\sigma}$ was replaced by $\sqrt{N(N - 2^3)}$ as a bias correction. It may be noted that, unlike the maximum likelihood (ML) estimator of σ, the MML estimator $\hat{\sigma}$ is always real and positive.

4. Properties of the estimators and hypotheses testing

The modified likelihood equations are asymptotically equivalent to the corresponding likelihood equations. Therefore, the MML estimators are asymptotically unbiased and efficient; see Bhattacharrya [4] and Şenoğlu [10]. The following results are instrumental for testing the null hypotheses;

$$H_{01} : \tau_i = 0 \ (i = 1, 2), H_{02} : \beta_j = 0 \ (j = 1, 2), H_{03} : \gamma_k = 0 \ (k = 1, 2),$$

$$H_{04} : (\tau\beta)_{ij} = 0 \ (i = 1, 2 \text{ and } j = 1, 2), \tag{4.1}$$

$$H_{05} : (\tau\gamma)_{ik} = 0 \ (i = 1, 2 \text{ and } k = 1, 2),$$

$$H_{06} : (\beta\gamma)_{jk} = 0 \ (j = 1, 2 \text{ and } k = 1, 2) \text{ and}$$

$$H_{07} : (\tau\beta\gamma)_{ijk} = 0 \ (i = 1, 2, j = 1, 2 \text{ and } k = 1, 2).$$

Lemma 1: The estimator $\hat{\tau}_i$ is an unbiased estimator of τ_i and is asymptotically normally distributed with variance $\sigma^2/4m(b+1)$.

Proof: The result follows from the fact that $\partial \ln L/\partial \tau_i$ assumes the form (see [6])

$$\frac{\partial \ln L}{\partial \tau_i} \cong \frac{\partial \ln L^*}{\partial \tau_i} = \frac{4m(b+1)}{\sigma^2}(\hat{\tau}_i - \tau_i)$$

with $E(\partial^r \ln L^*/\partial \tau_i^r) = 0$ for all $r \geq 3$, see Bartlett [2].

From the same argument given in the proof of Lemma 1, $\hat{\beta}_j$ and $\hat{\gamma}_k$ are unbiased estimators of β_j and γ_k, respectively, with variance $\sigma^2/4m(b+1)$ and they are asymptotically normally distributed.

Lemma 2: The estimator $\widehat{(\tau\beta)}_{ij}$ is an unbiased estimator of $(\tau\beta)_{ij}$ and is asymptotically normally distributed with variance $\sigma^2/2m(b+1)$.

Proof: Asymptotically, $\partial \ln L/\partial(\tau\beta)_{ij}$ assumes the form

$$\frac{\partial \ln L}{\partial(\tau\beta)_{ij}} \cong \frac{\partial \ln L^*}{\partial(\tau\beta)_{ij}} = \frac{2m(b+1)}{\sigma^2}[\widehat{(\tau\beta)}_{ij} - (\tau\beta)_{ij}].$$

From the same argument given in the proof of Lemma 2, $\widehat{(\tau\gamma)}_{ik}$ and $\widehat{(\beta\gamma)}_{jk}$ are unbiased estimators of $(\tau\gamma)_{ik}$ and $(\beta\gamma)_{jk}$, respectively, with variance $\sigma^2/2m(b+1)$ and they are asymptotically normally distributed.

Lemma 3: The estimator $\widehat{(\tau\beta\gamma)}_{ijk}$ is an unbiased estimator of $(\tau\beta\gamma)_{ijk}$ and is asymptotically normally distributed with variance $\sigma^2/m(b+1)$.

Proof: This follows from the fact that $\partial \ln L^*/\partial(\tau\beta\gamma)_{ijk}$ is asymptotically eqivalent to $\partial \ln L/\partial(\tau\beta\gamma)_{ijk}$ and assumes the form

$$\frac{\partial \ln L}{\partial(\tau\beta\gamma)_{ijk}} \cong \frac{\partial \ln L^*}{\partial(\tau\beta\gamma)_{ijk}} = \frac{m(b+1)}{\sigma^2}[\widehat{(\tau\beta\gamma)}_{ijk} - (\tau\beta\gamma)_{ijk}],$$

see Lemmas 1 and 2.

Lemma 4: Asymptotically, the MML estimators $\hat{\tau}_i$, $\hat{\beta}_j$, $\hat{\gamma}_k$, $\widehat{(\tau\beta)}_{ij}$, $\widehat{(\tau\gamma)}_{ik}$, $\widehat{(\beta\gamma)}_{jk}$, $\widehat{(\tau\beta\gamma)}_{ijk}$ and $\hat{\sigma}$ are independently distributed.

Proof: Asymptotic independence of τ_i and σ follows from the fact that $E(\partial^{r+s} \ln L^*/\partial \tau_i^r \partial \sigma^s)=0$ for all $r \geq 1$ and $s \geq 1$; see Bartlett [2]. The MML estimators $\hat{\beta}_j$, $\hat{\gamma}_k$, $\widehat{(\tau\beta)}_{ij}$, $\widehat{(\tau\gamma)}_{ik}$, $\widehat{(\beta\gamma)}_{jk}$, $\widehat{(\tau\beta\gamma)}_{ijk}$ are asymptotically independent of $\hat{\sigma}$ from the same reasons given for $\hat{\tau}_i$.

Lemma 5: $N\hat{\sigma}^2/\sigma^2$ is for large n (for known $\mu_{ijk.}$) referred to a chi-square distribution with $\upsilon = N = 2^k n$ degrees of freedom.
   Proof: Write

$$B_0 = (b+1)\sum_i \sum_j \sum_k \sum_l (y_{ijk(l)} - \mu_{ijk.})\Delta_l \quad \text{and}$$

$$C_0 = (b+1)\sum_i \sum_j \sum_k \sum_l (y_{ijk(l)} - \mu_{ijk.})^2 \beta_l.$$

Realizing that $B_0 \big/ \sqrt{NC_0} \cong 0$ asymptotically and $\partial \ln L/\partial \sigma$ assumes the form

$$\frac{\partial \ln L}{\partial \sigma} \cong \frac{\partial \ln L^*}{\partial \sigma} = \frac{N}{\sigma^3}\left(\frac{C_0}{N} - \sigma^2\right).$$

For testing the hypotheses given in (4.1), we define the following test statistics based on the MML estimators

$$F_A^* = \frac{4m(b+1)\sum_i \hat{\tau}_i^2}{\hat{\sigma}^2}, \quad F_B^* = \frac{4m(b+1)\sum_j \hat{\beta}_j^2}{\hat{\sigma}^2}, \quad F_C^* = \frac{4m(b+1)\sum_k \hat{\gamma}_k^2}{\hat{\sigma}^2}$$

$$F_{AB}^* = \frac{2m(b+1)\sum_i \sum_j (\widehat{\tau\beta})_{ij}^2}{\hat{\sigma}^2}, \quad F_{AC}^* = \frac{2m(b+1)\sum_i \sum_k (\widehat{\tau\gamma})_{ik}^2}{\hat{\sigma}^2} \qquad (4.2)$$

$$F_{BC}^* = \frac{2m(b+1)\sum_j \sum_k (\widehat{\beta\gamma})_{jk}^2}{\hat{\sigma}^2}, \quad F_{ABC}^* = \frac{m(b+1)\sum_i \sum_j \sum_k (\widehat{\tau\beta\gamma})_{ijk}^2}{\hat{\sigma}^2}$$

respectively.

   Asymptotically, the null distributions of the test statistics in (4.2) are referred to a central F distribution with degrees of freedom (ν_1, ν_8), (ν_2, ν_8), (ν_3, ν_8), (ν_4, ν_8), (ν_5, ν_8), (ν_6, ν_8) and (ν_7, ν_8), respectively:

$$\nu_1 = 1, \ \nu_2 = 1, \ \nu_3 = 1, \ \nu_4 = 1, \ \nu_5 = 1, \ \nu_6 = 1, \ \nu_7 = 1 \text{ and } \nu_8 = 2^3(n-1).$$

   To have an idea how accurate these central F approximations are, we simulated the probabilities

$$P_1 = prob\left[F_A^* \geq F_{0.05}(\nu_1, \nu_8) \mid H_{01}\right], \quad P_2 = prob\left[F_{AB}^* \geq F_{0.05}(\nu_4, \nu_8) \mid H_{04}\right] \tag{4.3}$$

and

$$P_3 = prob\left[F_{ABC}^* \geq F_{0.05}(\nu_7, \nu_8) \mid H_{07}\right],$$

from 10,000 Monte Carlo runs. The values are given in Table 1. It should be noted that all of the main effects and the two-factor interactions have similar power properties, therefore, we will consider only one of the main effects and one of the two-factor interactions for conciseness.

TABLE 1. Simulated values of the probabilities.

		b					
		0.5	1.0	2.0	3.0	4.0	6.0
$n = 4$	F_A^*	0.050	0.048	0.048	0.046	0.045	0.044
	F_{AB}^*	0.046	0.047	0.048	0.046	0.046	0.045
	F_{ABC}^*	0.044	0.040	0.042	0.042	0.041	0.040
$n = 5$	F_A^*	0.046	0.045	0.048	0.048	0.048	0.047
	F_{AB}^*	0.049	0.046	0.046	0.043	0.042	0.043
	F_{ABC}^*	0.048	0.046	0.046	0.045	0.046	0.044
$n = 6$	F_A^*	0.050	0.049	0.049	0.045	0.044	0.043
	F_{AB}^*	0.051	0.046	0.047	0.046	0.047	0.045
	F_{ABC}^*	0.054	0.053	0.050	0.050	0.049	0.048

The central F-distribution gives remarkably accurate approximations even for small n.

The traditional F statistics based on LS estimators are given by

$$F_A = \frac{[(a-1)(b+1)(c+1)]^2}{2^3 n \tilde{\sigma}^2}, \quad F_B = \frac{[(a+1)(b-1)(c+1)]^2}{2^3 n \tilde{\sigma}^2},$$

$$F_C = \frac{[(a+1)(b+1)(c-1)]^2}{2^3 n \tilde{\sigma}^2}, \quad F_{AB} = \frac{[(a-1)(b-1)(c+1)]^2}{2^3 n \tilde{\sigma}^2}, \qquad (4.4)$$

$$F_{AC} = \frac{[(a-1)(b+1)(c-1)]^2}{2^3 n \tilde{\sigma}^2}, \quad F_{BC} = \frac{[(a+1)(b-1)(c-1)]^2}{2^3 n \tilde{\sigma}^2},$$

$$F_{ABC} = \frac{[(a-1)(b-1)(c-1)]^2}{2^3 n \tilde{\sigma}^2} \quad \text{and} \quad \tilde{\sigma}^2 = \sum_i \sum_j \sum_k \sum_l (y_{ijkl} - \bar{y}_{ijk.})^2/(N - 2^3).$$

It should be noted that treatments are the combinations of the factor levels and are represented by the letters (1), a, b, c, ab, ac, bc and abc. However, in formula (4.4) they represent the totals of n observations in each treatment, see Montgomery [7] for more information.

The distributions of F_A, F_B, F_C, F_{AB}, F_{AC}, F_{BC} and F_{ABC} are central or noncentral F depending on whether H_{0i} ($i =$1, 2, 3, 4, 5, 6, 7) are true or not.

Given in Table 2 are the simulated values of the type I error and power of the F_{ABC} and F^*_{ABC} tests; σ was taken to be equal to 1 without loss of generality, and presumed value of the type I error is 0.050. Simulation results show that the power of the F_{ABC} test is considerably lower than that of the F^*_{ABC} test.

TABLE 2. Values of the power of F_{ABC} and F^*_{ABC} tests; n=4.

b	$d =$	0.00	0.15	0.30	0.45	0.60	0.75	0.90	1.05	1.20	1.35
0.5	F_{ABC}	0.043	0.06	0.09	0.16	0.26	0.38	0.50	0.63	0.73	0.82
	F^*_{ABC}	0.044	0.06	0.10	0.19	0.29	0.42	0.56	0.69	0.80	0.88
1.0	F_{ABC}	0.043	0.07	0.15	0.28	0.45	0.63	0.78	0.89	0.89	0.98
	F^*_{ABC}	0.040	0.07	0.15	0.28	0.45	0.63	0.78	0.89	0.89	0.98
2.0	F_{ABC}	0.041	0.08	0.20	0.38	0.59	0.78	0.90	0.97	0.99	1.00
	F^*_{ABC}	0.042	0.08	0.20	0.40	0.61	0.81	0.92	0.97	0.99	1.00
4.0	F_{ABC}	0.042	0.09	0.23	0.44	0.67	0.85	0.94	0.98	1.00	1.00
	F^*_{ABC}	0.041	0.09	0.24	0.48	0.72	0.89	0.97	0.99	1.00	1.00
6.0	F_{ABC}	0.041	0.09	0.24	0.46	0.70	0.86	0.95	0.99	1.00	1.00
	F^*_{ABC}	0.040	0.09	0.26	0.51	0.76	0.91	0.98	1.00	1.00	1.00

Robustness: The true value of shape parameter b may, in practice, be somewhat different from the one assumed. In this section, we study the robustness of the test statistics based on the MML and the LS estimators given in (4.2) and (4.4), respectively, to understand how robust the test statistics are with respect to plausible deviations from an assumed model.

The value of b is assumed to be 2 for illustration in $GL(b,\sigma)$. In fact, any other value of b can be chosen with similar results. The model $GL(2,\sigma)$ will be called population model. The alternatives to this model will be called sample models. Out of a large number of plausible sample models, we choose the following sample models;

(1) $b=1.5$, (2) $b=3.0$;
(3) Dixon's outlier model: (n-1) observations come from $GL(2,\sigma)$ but one observation (we do not know which one) comes from $GL(2,2\sigma)$;
(4) Mixture model: $0.90GL(2,\sigma)+0.10GL(2,2\sigma)$;
(5) Contamination model: $0.90GL(2,\sigma)+0.10$Uniform(-1,1).

Given in Table 3 are the values of the power of the F_{ABC} and F^*_{ABC} statistics. It is clear that F^*_{ABC} test has higher power than the traditional F_{ABC} test based on LS estimators. Therefore, it is remarkably robust to deviations from an assumed $GL(b,\sigma)$.

TABLE 3. Values of the power for alternatives to GL(2, sigma); n=4; alfa=0.050.

	$d =$	0.0	0.2	0.4	0.6	0.8	1.0	1.2
True Model	F^*_{ABC}	0.042	0.11	0.33	0.61	0.85	0.96	0.99
	F_{ABC}	0.042	0.11	0.31	0.59	0.83	0.95	0.99
Model (1)	F^*_{ABC}	0.043	0.10	0.29	0.55	0.80	0.93	0.98
	F_{ABC}	0.042	0.10	0.28	0.54	0.79	0.93	0.98
Model (2)	F^*_{ABC}	0.040	0.11	0.37	0.67	0.89	0.98	1.00
	F_{ABC}	0.041	0.12	0.35	0.64	0.87	0.97	0.99
Model (3)	F^*_{ABC}	0.042	0.10	0.30	0.58	0.83	0.95	0.99
	F_{ABC}	0.043	0.09	0.28	0.55	0.80	0.93	0.98
Model (4)	F^*_{ABC}	0.038	0.10	0.31	0.60	0.84	0.96	0.99
	F_{ABC}	0.039	0.10	0.29	0.57	0.81	0.94	0.99
Model (5)	F^*_{ABC}	0.041	0.13	0.35	0.63	0.86	0.96	0.99
	F_{ABC}	0.043	0.13	0.34	0.61	0.84	0.95	0.99

5. GENERALIZATION TO K-FACTOR CASES

Suppose that we have K factors, each at two levels, i. e. 2^k factorial design. Sum of squares for the main effects and the interactions (two-factor, three-factor,...., k-factor etc.) are given by

$$SS_A = 2^{k-1}m(b+1)\sum_a \hat{A}_a^2, \quad SS_B = 2^{k-1}m(b+1)\sum_b \hat{B}_b^2, \ldots\ldots,$$

$$SS_K = 2^{k-1}m(b+1)\sum_k \hat{K}_k^2,$$

$$SS_{AB} = 2^{k-2}m(b+1)\sum_a \sum_b (\widehat{AB})_{ab}^2,$$

$$SS_{AC} = 2^{k-2}m(b+1)\sum_a \sum_c (\widehat{AC})_{ac}^2, \ldots\ldots,$$

$$SS_{JC} = 2^{k-2}m(b+1)\sum_j \sum_c (\widehat{JC})_{jc}^2,$$

$$SS_{ABC} = 2^{k-3}m(b+1)\sum_a \sum_b \sum_c (\widehat{ABC})_{abc}^2,$$

$$SS_{ABD} = 2^{k-3}m(b+1)\sum_a \sum_b \sum_d (\widehat{ABD})_{abd}^2, ...,$$

$$SS_{IJK} = 2^{k-3}m(b+1)\sum_i \sum_j \sum_k (\widehat{IJK})_{ijk}^2$$

and

$$SS_{ABC......K} = m(b+1)\sum_a \sum_b .....\sum_k (\widehat{ABC...K})^2_{abc.....k}.$$

The mean square error is found by taking square of $\hat{\sigma} = \frac{B + \sqrt{B^2 + 4NC}}{2\sqrt{N(N-2^k)}}$. Here,

$$B = (b+1)\sum_a \sum_b ....\sum_k \sum_l \left(y_{abc....k(l)} - \hat{\mu}_{abc....k.}\right)\Delta_l \ \text{and}$$

$$C = (b+1)\sum_a \sum_b ....\sum_k \sum_l \left(y_{abc....k(l)} - \hat{\mu}_{abc....k.}\right)^2 \beta_l.$$

6. Conclusions

In this study, we extend the results of Şenoğlu [10] to the case where the underlying distribution of error terms is generalized logistic. We obtained the estimators of the model parameters by using the methodology known as modified maximum likelihood and proposed new test statistics based on these estimators. Simulation results reveal that our test statistics have higher power and are more robust than the traditional tests statistics based on LS estimators.

2^k FAKTORİYEL TASARIMDA DAYANIKLI TAHMİN VE HİPOTEZ TESTİ

ÖZET: Bu çalışma, hata terimlerinin Weibull dağılıma sahip olması durumunda sağlam 2^k faktöriyel tasarımlar isimli çalışmamın genişletilmiş bir halidir. Bu makalede, Uyarlanmış En Çok Olabilirlik (UEÇO) metodolojisi kullanılarak hata terimlerinin Genelleştirilmiş Lojistik dağılıma sahip olması durumunda 2^k faktöriyel tasarımdaki parametreler için sağlam ve etkin tahmin ediciler bulunmuş, ana etki ve etkileşimleri test etmek için UEÇO tahmin edicilerine dayanan yeni test istatistikleri önerilmiştir. Önerilen test istatistiklerinin, En Küçük Kareler (EKK) tahmin edicilerine dayanan test istatistiklerinden daha sağlam ve güçlü oldukları gösterilmiştir.

References

[1] N. Balakrishnan, W. Y. Leung, Means, variances and covariances of order statistics, BLUE for the Type I generalized logistic distribution and some applications. Commun. Stat.-Simula. 17 (1988) 51-84.

[2] M. S. Bartlett, Approximate confidence intervals, Biometrika 40 (1953) 12-19.

[3] J. Berkson, Why I prefer logits to probits, Biometrics 7 (1951) 327-339.

[4] G. K. Bhattacharrya, The asymptotics of maximum likelihood and related estimators based on type II censored data, Journal of the American Statistical Association 80 (1985) 398-404.

[5] R. A. Fisher, The Design of Experiments, Oliver & Boyd, Edinburgh, 1935.

[6] M. G. Kendall, A. Stuart, The Advanced Theory of Statistics, Charles Griffin, London, 1979.

[7] D. C. Montgomery, Design and Analysis of Experiments, John Wiley & Sons, New York, 1984.

[8] S. Puthenpura, N. K. Sinha, Modified maximum likelihood method for the robust estimation of system parameters from very noisy data, Automatica 22 (1986) 231-235.

[9] R. L. Smith, Maximum likelihood estimation in a class of nonregular cases, Biometrika 72 (1985) 67-90.

[10] B. Şenoğlu, Robust 2^k factorial design with Weibull error distributions, Journal of Applied Statistics 32 (10) (2005) 1051-1066.

[11] B. Şenoğlu, M. L. Tiku, Analysis of variance in experimental design with nonnormal error distributions, Communications in Statistics-Theory and Methods 30 (2001) 1335-1352.

[12] M .L. Tiku, Estimating the mean and standard deviation from censored normal samples, Biometrika 54 (1967) 155-165.

[13] D. C. Vaughan, The Generalized Secant Hyperbolic distribution and its properties, Communications in Statistics-Theory and Methods 31 (2002) 219-238.

[14] F. Yates, Design and Analysis of Factorial Experiments, Tech. Comm. No. 35, Imperial Bureau of Soil Sciences, London, 1937.

Current address: Ankara University, Faculty of Sciences, Department of Statistics, 06100 Tandoğan-Ankara, Turkey

E-mail address: senoglu@science.ankara.edu.tr