Sigma J Eng Nat Sci, Vol. 43, No. 1, pp. 73–87, February, 2025

Sigma Journal of Engineering and Natural Sciences
Web page info: https://sigma.yildiz.edu.tr
DOI: 10.14744/sigma.2025.00006

Research Article

Testing equality of two-parameter exponentially distributed mean lifetimes under unequal failure rates

Mustafa ÇAVUS[1,*], Berna YAZICI[1]

[1]*Department of Statistics, Eskisehir Technical University, Eskisehir, 26555, Türkiye*

ARTICLE INFO

Article history
Received: 20 September 2023
Revised: 13 November 2023
Accepted: 15 January 2024

Keywords:
Computational Approach; Non-Normal ANOVA; Penalized Power; Unequal Scales

ABSTRACT

Testing the equality of means of several skewed populations, particularly in the presence of nuisance parameters, is a central challenge in statistics. While various tests have been proposed for such as log-normal, inverse-normal, and exponential distributions leveraging methods like generalized p-value, parametric bootstrap, and the fiducial approach, there remains a notable gap in the literature, the absence of a computational approach method-based test for the two-parameter exponential distribution. Such a method is essential for achieving robust results in small sample sizes while considering power and Type I error probability. In response to this gap, our paper introduces and implements novel computational approach tests embedded in the doex package in R. Our focus is on assessing the equality of means for several skewed populations following a two-parameter exponential distribution. We conduct a comprehensive comparison of our proposed tests against existing alternatives, evaluating their penalized power and Type I error probability. Notably, our computational approach tests exhibit superior performance, particularly in cases involving small samples and balanced designs. Furthermore, to illustrate the practical relevance of our proposed tests, we present a real-world application using authentic data. This empirical demonstration serves to underscore the efficacy and applicability of our novel computational approach tests in real-world scenarios.

Cite this article as: Çavus M, Yazıcı B. Testing equality of two-parameter exponentially distributed mean lifetimes under unequal failure rates. Sigma J Eng Nat Sci 2025;43(1):73–87.

INTRODUCTION

Testing equality of means in the presence of nuisance parameters is a well-known Behrens-Fisher-type problem in statistics. There are many methods to solve this problem, such as generalized p-value, the parametric bootstrap, the fiducial approach, and the computational approach method. These methods are applied to many different Behrens-Fisher-type problems. Weerahandi [1] proposed the Generalized F-test, Krishnamoorthy et al. [2] proposed a parametric bootstrap test, Li et al. [3] proposed a fiducial approach test and Gokpinar and Gokpinar [4] proposed a computational approach test for testing equality of several means of normal populations under unequal variances. For comparing several log-normal means, Gokpinar and Gokpinar [5] and Jafari and Abdollahnezhad [6] proposed

***Corresponding author.**
*E-mail address: mustafacavus@eskisehir.edu.tr

This paper was recommended for publication in revised form by Editor-in-Chief Ahmet Selim Dalkilic

Published by Yıldız Technical University Press, İstanbul, Turkey

tests based on the computational approach method. It is seen that this method has been used frequently in recent years. The major advantage of the computational approach over the alternatives is to be able to derive powerful tests in small samples without the need for knowledge about the sample distribution. The limitation of this method is discussed in [7].

The two-parameter exponential distribution is used in many real-life problems such as modeling extreme rainfalls, the lifetime of a component, the service time of an agent, and so on. There are some procedures improved for the two-parameter exponential distribution. Chen [8] proposed a range statistic for comparing location parameters of two-parameter exponential distributions. Singh [9] derived a likelihood ratio test for testing the equality of location parameters of two-parameter exponential distributions based on Type II censored samples under unknown scales. Kambo and Awad [10] proposed a test statistic based on doubly censored samples to test the equality of location parameters of k exponential distributions when the scale parameter is unknown. Hsieh [11] proposed an exact test for comparing location parameters simultaneously of several two-parameter exponential distributions under unequal scale parameters with unknown scale parameters. Vaughan and Tiku [12] extended the test developed by Tiku and Vaughan [13] for k - 2 populations for testing the equality of location parameters of two-parameter exponential populations from censored samples. Ananda and Weerahandi [14] proposed a testing procedure based on generalized p-values for testing the difference of two exponential means. Wu [15] proposed a one-stage multiple comparison procedure for comparing k - 1 treatment exponential mean lifetimes with the control based on doubly censored samples under unequal scales. Malekzadeh and Jafari [16] proposed some procedures based on generalized p-values, parametric bootstrap, and fiducial approach by using Cochran type test statistics for testing the means of several two-parameter exponential distributions under progressively Type II censoring. In the testing equality of means of two-parameter exponential distributions, the scale parameter is a nuisance parameter when it is unknown or unequal. Therefore, the considered problem turns into a Behrens-Fisher-type problem. The major contribution of this paper is proposing the computational approach tests for testing the equality of two-parameter exponentially distributed population means under unequal scale parameters.

The following section introduces the alternative tests and proposed computational approach tests. In Section 2, the alternative tests are introduced. Our proposed CATs are introduced in Section 3. The performance of the proposed tests with the alternatives is investigated in terms of penalized power and Type I error probability in Section 4. An illustrative example is given in Section 5. The results are discussed in the last section.

TESTING EQUALITY OF SEVERAL TWO-PARAMETER EXPONENTIAL MEANS UNDER UNEQUAL SCALES

In this section, the generalized p-value, parametric bootstrap, and fiducial approach tests are given for testing the equality of two-parameter exponential distributed populations' means. The probability density function of the two-parameter exponential distribution is as follows:

$$f(x;\ a,b) = \frac{1}{a} exp\left(-\frac{x-b}{a}\right),\ \ x > b,\ \ a > 0 \tag{1}$$

where a is the scale and b is the location parameter. The interested hypothesis is in the following for testing the equality of means of the exponentially distributed populations.

$$H_0: \mu_1 = \mu_2 = \ldots = \mu_k \tag{2}$$

where μ_i is the mean of the i.th population. Johnson and Kotz [17] obtained the maximum likelihood estimate of the parameters as in the following equations [18, 19]:

$$\hat{a} = S/(n-1),\quad \hat{b} = X_{(1)} \tag{3}$$

where $X_{(1)}{=}min(X_1, X_2, \ldots, X_n)$ and $S{=}\sum_{j=1}^{n}(X_j{-}X_{(1)})$. Cochran [20] type test statistics are used for Behrens-Fisher problems. Here, it is modified for testing the equality of two-parameter exponential distributed means under unequal scale parameters.

$$T_t = \sum_{i=1}^{k} \frac{n_i \mu_i^2}{S_i^2} - \frac{\left(\sum_{i=1}^{k} \frac{n_i \mu_i^2}{S_i^2}\right)}{\sum_{i=1}^{k} \frac{n_i}{S_i^2}} \tag{4}$$

where μ_i is the mean estimate and S_i is the scale estimate. The uniformly minimum variance unbiased estimator of $\hat{\mu}_i = \hat{a}_i + \hat{b}_i$ can be shown as in (5).

$$\hat{\mu}_i = X_{i(1)} + \frac{n_i - 1}{n_i}\hat{S}_i \sim N(\mu_i, {a_i}^2/n_i) \tag{5}$$

where $X_{i(1)}$ is the first order statistics and $\hat{S}_i = \sum_{j=1}^{n}(X_{ij} - X_{i(1)})$ is the scale estimator of the i.th population. Under the null hypothesis, T_t is approximately chi-square distributed with k - 1 degrees of freedom. T_t is used for the rejection rule as a critical value of the generalized p-value, parametric bootstrap, fiducial approach, and proposed computational approach tests in the following sections.

Generalized p-value (GP) Test

Tsui and Weerahandi [21] introduced the concept of generalized p-value can be used to derive the test statistics in the presence of nuisance parameters. Many researchers used this method to derive test statistics for several distributions. In this method, firstly sufficient statistics of parameters of the related distribution are obtained. Using the sufficient statistics of the two-parameter exponential distribution, Malekzadeh and Jafari [16] proposed the GP test for the testing hypothesis in (2) by following these steps: (i) R_i can be obtained independently from the nuisance parameter and, (ii) since the observed λ_i values are independent of the nuisance parameter θ_i, generalized pivot variable can be estimated.

$$R_i = x_{i(1)} + (n_i - 1)s_i(2n_i - V_i/n_iU_i) \tag{6}$$

where U_i and V_i are independent random samples with $U_i \sim \chi^2_{(2n_i-2)}$ and $V_i \sim \chi^2_{(2)}$. For a given observed values of $[X_{i(1)}, S_i]$ as $[x_{i(1)}, s_i]$, the expected value and the variance of the generalized pivot variable R_i are obtained as follows:

$$\mu_{Ri} = x_{i(1)} + \frac{(n_i-1)^2 s_i}{n_i^2 - 2n_i}, \quad \sigma_{Ri}^2 = \frac{(n_i-1)^4 s_i^2}{n_i^2(n_i-2)^2}\left(\frac{1}{n_i-3}\right) \tag{7}$$

The generalized p-value test statistic is obtained as in (8) using the expected value and variance of the generalized pivot variable R_i.

$$T_{GP} = \sum_{i=1}^{k} \frac{\mu_{Ri}^2}{\sigma_{Ri}^2} - \frac{\left(\sum_{i=1}^{k} \frac{\mu_{Ri}}{\sigma_{Ri}^2}\right)^2}{\sum_{i=1}^{k} \frac{1}{\sigma_{Ri}^2}} \tag{8}$$

The rejection rule is *H0* in (2) rejected when $T_{GP} \geq T_t$. The p-value of the GP test is computed at least *10,000* for Monte-Carlo runs $p_{GP} = P(T_{GP} \geq T_t)$. The null hypothesis is rejected when $p_{GP} < \alpha_0$.

Parametric Bootstrap (PB) Test

Krishnamoorthy et al. [2] used the parametric bootstrap approach for testing the equality of several population means under unequal variances. Malekzadeh and Jafari [16] proposed the PB test for the testing hypothesis in (2).

$$T_{PB} = \sum_{i=1}^{k} \frac{n_i\mu_{Bi}^2}{S_i^2} - \frac{\left(\sum_{i=1}^{k} \frac{n_i\mu_{Bi}}{S_i^2}\right)^2}{\sum_{i=1}^{k} \frac{n_i}{S_i^2}} \tag{9}$$

where $V_i \sim \chi^2_{(2)}$ and $U_i \sim \chi^2_{(2n_i-2)}$ are independent random samples, $\mu_B = (s_i/2n_i)(V_i + U_i)$ and $S_{Bi} = s_iU_i/(2n_i-2)$. The p-value of the PB test is computed with *10,000* Monte-Carlo runs as in (10).

$$p_{PB} = P(T_{PB} \geq T_t) \tag{10}$$

The null hypothesis is rejected when $p_{PB} < \alpha_0$.

Fiducial Approach (FA) Test

Li et al. [4] proposed to use of the fiducial approach for testing equality of several population means under unequal variances. Malekzadeh and Jafari [16] proposed the PB test for the testing hypothesis in (2). Let $U_i \sim \chi^2_{(2n_i-2)}$ and $V_i \sim \chi^2_{(2)}$ are independent random samples. $X_{i(1)}$, and S_i functions can be rewritten as random samples:

$$S_i = \frac{\mu_i U_i}{2(n_i-1)}, \quad X_{i(1)} = \frac{\mu_i V_i}{2n_i} + \alpha_i \tag{11}$$

Parameter estimations are obtained as follows by using the observed values of $[X_{i(1)}, S_i]$:

$$a_i = x_{i(1)} - \frac{(n_i-1)s_iv_i}{n_iu_i}, \quad \mu_i = \frac{2(n_i-1)s_i}{u_ş}. \tag{12}$$

The test statistic T_{FA} is obtained as in (13).

$$T_{FA} = \sum_{i=1}^{k} \frac{f_i}{l_in_i^2u_i^2} - \frac{\left(\sum_{i=1}^{k} \frac{f_i}{l_is_in_iu_i}\right)^1}{\sum_{i=1}^{k} w_i} \tag{13}$$

where $f_i = (n_i - 1)(u_iv_i - 2n_iu_i)$. The p-value of the FA test is computed with 10,000 Monte-Carlo runs as in (14).

$$p_{FA} = P(T_{FA} \geq T_t) \tag{14}$$

The null hypothesis is rejected when $p_{FA} < \alpha_0$.

PROPOSED COMPUTATIONAL APPROACH TESTS (CATs)

Pal et al. [22] proposed the computational approach method which is a type of parametric bootstrap method. The CAT method based on simulation and numerical computations uses maximum likelihood estimates and does not require knowledge of any sampling distribution. It can be used easily, because of the development of calculation technology. Let δ is the nuisance parameter of the parameter space $\zeta = (\theta, \delta)$ and $T = T(X; x, \zeta)$ is considered test statistic in testing the hypothesis H_0: $\theta = \theta_0$ versus H_A: $\theta \neq \theta_0$. The p-value of the CAT is calculated by using the Algorithm 1.

To improve the test statistics, the RML estimators are obtained in the following steps. The likelihood function *T* (.) of a two-parameter exponential distribution is as follows:

Algorithm 1. p-value calculation of the CAT

1: Calculate the observed values of the interested parameter θ_i and nuisance parameter δ_i for *k* samples.
2: Calculate the observed value of the test statistics T^*_{CAT} using the estimator in Step 1.
3: Find the restricted maximum likelihood (RML) estimator of the parameters for *k* samples under the true H_0.
4: **for** $j \leftarrow 1$ to *M* **do**
5:     **for** i← 1 to n_i **do**
6:         **for** $j \leftarrow 1$ to *k* **do**
7:             Generate random samples from $X_{ij} \sim f(\hat{\theta}_{RML}, \hat{\delta}_{RML})$ for *k* samples.
8:         **end for**
9:     **end for**
10:     Calculate the observed values of the parameters for generated samples.
11:     Calculate the observed value of test statistics $T^{(m)}_{CAT}$ using the estimators in Step 8.
12: **end for**
13: Calculate the p-value of the CAT: $p = \sum_{l=1}^{M} I(t^{(m)}_{CAT} > t^*_{CAT})/M$

$$
\begin{aligned}
L(a,b \mid x_1,x_2,\dots,x_n) &= f(x_1 \mid a,b)\, f(x_2 \mid a,b) \dots f(x_n \mid a,b) \\
&= \prod_{i=1}^{n} f(x_i \mid a,b) \\
&= \prod_{i=1}^{n} \frac{1}{a} exp\left(-\frac{x_i - b}{a}\right) \\
&= \frac{1}{a^n} \exp\left[-\frac{1}{a}\sum_{i=1}^{n}(x_i - b)\right]
\end{aligned} \tag{15}
$$

Here, to obtain RML estimators, under the true H_0 hypothesis, the *b* parameter can be expressed $b = \mu - a$ in terms of the nuisance parameter.

$$
L(a,\mu - a \mid x_1,x_2,\dots,x_n) = \frac{1}{a^n}\exp\left[-\frac{1}{a}\sum_{i=1}^{n}(x_i - \mu + a)\right] \tag{16}
$$

The log-likelihood function of the two-parameter exponential distribution is as follows:

$$
\begin{aligned}
L(a,\mu - a \mid x_1,x_2,\dots,x_n) &= \prod_{i=1}^{n} \ln f(x_i \mid a, \mu - a) \\
&= -\frac{1}{a}\sum_{i=1}^{n} x_i + \frac{n\mu}{n} - \frac{na}{a} - n\, lna \\
&= -\frac{1}{a}\sum_{i=1}^{n} x_i + \frac{n\mu}{n} - n - n\, lna
\end{aligned} \tag{17}
$$

The maximum likelihood estimator of the parameters is obtained by the derivative of the log-likelihood function as follows:

$$
\frac{d\, lnL(a,\mu - a)}{da} = \frac{\sum_{i=1}^{n}(x_i - \mu)}{a^2} - \frac{n}{a} = 0 \tag{18}
$$

Algorithm 2. p-value calculation of the proposed CATs

1: Calculate the observed values of μ_i and a_i for *k* samples.
2: Calculate the observed value of the test statistics T_t using the estimator in Step 1.
3: Find the restricted maximum likelihood (RML) estimator of the parameters for *k* samples under the true H_0.
4: **for** $j \leftarrow 1$ to *M* **do**
5:     **for** i← 1 to n_i **do**
6:         **for** $j \leftarrow 1$ to *k* **do**
7:             Generate random samples from $X_{ij} \sim f(\hat{a}_{i(RML)}, \hat{\mu}_{i(RML)} - \hat{a}_{i(RML)})$ for *k* samples.
8:         **end for**
9:     **end for**
10:     Calculate the observed values of the μ_i and a_i for generated samples.
11:     Calculate the observed value of test statistics $T^{(m)}_t$ using the estimators in Step 8.
12: **end for**
13: Calculate the p-value of the proposed CAT: $p = \sum_{l=1}^{M} I(t^{(m)}_t > t_t)/M$

The restricted maximum likelihood estimator of the scale parameter is obtained in (19).

$$\hat{a}_{RML} = \mu - x_{(1)} \tag{19}$$

The $\hat{a}_{RML}$ estimator obtained here is used to calculate the critical values of CATs. The maximum likelihood estimations of the two-parameter exponential distribution are biased. Zheng [23] obtained the penalized maximum likelihood estimators, and the estimations of the parameters have lower MSEs obtained. In this paper, there are four different CATs are proposed for testing the equality of two-parameter exponential distributed populations' means by using the combinations of the mean estimators. The CAT1, CAT2, CAT3, and CAT4 are introduced in the following subsections.

Proposed Computational Approach Test 1 (CAT1)

The maximum likelihood estimator of the mean $\hat{\mu}_i = \hat{a}_i + \hat{b}_i$ is used in CAT1. It is also labeled as reference tests, and the null hypothesis is rejected when $p_{CAT1} < \alpha_0$, which is calculated as in Algorithm 2.

Proposed Computational Approach Test 2 (CAT2)

The maximum likelihood estimation of the mean with the correction of unbiasedness $\hat{\mu}_i = \hat{a}_i + \frac{n_i - 1}{n_i} x_{(1)}$ as in [16] is used in CAT2. The null hypothesis is rejected when $p_{CAT2} < \alpha_0$.

Proposed Computational Approach Test 3 (CAT3)

The penalized maximum likelihood estimation of the mean with the correction of unbiasedness $\hat{\mu}_i = \hat{a}_i + \frac{n_i - 1}{n_i} \hat{b}_{RML}$ as in [16] is used in CAT3. The null hypothesis is rejected when $p_{CAT3} < \alpha_0$.

Proposed Computational Approach Test 4 (CAT4)

The penalized maximum likelihood estimation of the mean $\hat{\mu}_i = \hat{a}_i + \hat{b}_{RML}$ is used in CAT4. The null hypothesis is rejected when $p_{CAT4} < \alpha_0$. The performance of the proposed CATs over the GPD, PB, and FY tests is investigated by Monte-Carlo simulation studies in the next section.

MONTE-CARLO SIMULATION STUDY

The performance of the proposed CATs is investigated over the alternatives in terms of penalized power and Type I error probability when the nominal level of the test is taken $\alpha = 0.05$ under different sample sizes and scale parameters in this section. We provide comprehensive simulation study results. It is known that Monte-Carlo simulation studies are used to compare the performance of the tests in terms of power and Type I error probability. However, any comparison of the powers is invalid when Type I error probabilities are different. Cavus et al. [24] proposed the penalized power approach to compare the power of the tests when Type I error probabilities are different.

$$\gamma_i = \frac{1 - \beta_i}{\sqrt{1 + \left|1 - \frac{\alpha_i}{\alpha_0}\right|}} \tag{20}$$

where β is the Type II error rate, α_i is the Type I error of the test and α_0 is the nominal level. Penalized power adjusts the power function with the square root of the percentile deviation between Type I error probability and the nominal level. Thus, penalized power is used to compare the power of the tests in the simulation studies. The simulations are performed for balanced and unbalanced designs with doex package implemented in R [25, 26], and the results are based on 10,000 Monte-Carlo replications. The results of the simulations are given in the following subsections.

Type I Error Probability Results

The proposed CATs and the alternatives in the literature to control Type 1 error probability relative to the nominal level are investigated. In the simulation study, the scale parameter is a_i, the location parameter is b_i, the sample size is n_i, and the number of populations is k taken as configuration factors. In addition, balanced and unbalanced designs are considered, and by increasing the differences between the sample size in the designs; the effect of design type on performance is investigated. The nominal level $\alpha_0 = 0.05$ and location parameters are fixed as $b_i = 1,1,1$. The Type I error probability of the tests $k = 3$ is given in Table 1. In the following tables, GP refers generalized p-value test, PB refers parametric bootstrap test, FA refers fiducial approach test, and CATs refer to computational approach tests.

According to the results in Table 1, CAT1, CAT2, and CAT4 can control the Type I error probability in small samples, while the PB test shows similar performance in medium and large samples. On the other hand, GPV and FA tests can only control the Type I error probability in large samples. When the properties of the tests to control Type I error probability are compared, GP, PB, FA, and CAT3 tests are conservative, and CAT1, CAT2, and CAT4 are mostly liberal but conservative in some cases. The results of the Type I error probability of the tests $k = 4$ are given in Table 2.

The increasing of k from 3 to 4, decreases the Type I error probability of the tests except CAT4. In the general framework, GPV, PB, FA, and CAT3 tests are conservative, while CAT1 and CAT2 tests are liberal in terms of Type I error probability. CAT4 test has a Type I error probability higher than the nominal level except for small samples. In medium samples, the Type I error probability of the GP and FA tests is far from the nominal level. In large samples, the FA test, as well as balanced designs, in addition to the GP and PB test, Type 1 error probability is far from the nominal level. While an unbalanced design does not negatively affect the performance of GP, PB, and FA tests, CATs seem to affect increasing the Type 1 error probability. This causes the Type 1 error probability of the CATs to exceed the nominal level. In balanced designs, CATs generally perform better, in all cases the Type 1 error probability of the CAT4 is very

Table 1. Type I error probability for $k = 3$

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
10,10,10	2,2,2	1, 1, 1	0.006	0.008	0.001	0.046	0.042	0.012	0.038
	5,5,5		0.006	0.010	0.001	0.055	0.059	0.017	0.041
8,10,12	2,2,2		0.006	0.012	0.002	0.065	0.072	0.031	0.039
	5,5,5		0.008	0.013	0.002	0.053	0.056	0.039	0.053
5,10,15	2,2,2		0.014	0.043	0.007	0.053	0.055	0.041	0.054
	5,5,5		0.012	0.043	0.007	0.055	0.057	0.044	0.043
30,30,30	2,2,2		0.026	0.040	0.020	0.055	0.057	0.049	0.055
	5,5,5		0.029	0.046	0.022	0.055	0.054	0.048	0.053
24,30,36	2,2,2		0.030	0.045	0.023	0.056	0.054	0.042	0.045
	5,5,5		0.029	0.045	0.022	0.057	0.056	0.019	0.045
15,30,45	2,2,2		0.033	0.048	0.024	0.055	0.059	0.017	0.041
	5,5,5		0.032	0.052	0.024	0.065	0.050	0.031	0.039
50,50,50	2,2,2		0.036	0.047	0.031	0.053	0.056	0.039	0.053
	5,5,5		0.034	0.045	0.030	0.053	0.055	0.041	0.054
40,50,60	2,2,2		0.035	0.049	0.031	0.055	0.057	0.044	0.043
	5,5,5		0.036	0.049	0.030	0.055	0.057	0.049	0.055
25,30,75	2,2,2		0.039	0.049	0.030	0.055	0.054	0.048	0.053
	5,5,5		0.040	0.051	0.033	0.051	0.048	0.041	0.041

Table 2. Type I error probability for $k = 4$

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
10,10,10,10	2,2,2,2	1, 1, 1, 1	0.008	0.011	0.001	0.071	0.067	0.020	0.058
	5,5,5,5		0.008	0.011	0.001	0.071	0.067	0.020	0.058
6,8,10,14	2,2,2,2		0.008	0.018	0.003	0.082	0.079	0.023	0.048
	5,5,5,5		0.008	0.018	0.003	0.082	0.079	0.023	0.048
4,8,12,16	2,2,2,2		0.009	0.037	0.005	0.088	0.089	0.026	0.042
	5,5,5,5		0.009	0.037	0.005	0.088	0.089	0.027	0.042
30,30,30,30	2,2,2,2		0.033	0.044	0.027	0.059	0.056	0.040	0.057
	5,5,5,5		0.033	0.044	0.027	0.059	0.056	0.040	0.057
18,24,30,42	2,2,2,2		0.028	0.039	0.018	0.050	0.052	0.039	0.044
	5,5,5,5		0.028	0.039	0.018	0.050	0.052	0.039	0.044
12,24,36,48	2,2,2,2		0.031	0.051	0.029	0.054	0.057	0.051	0.049
	5,5,5,5		0.031	0.051	0.029	0.054	0.057	0.051	0.049
50,50,50, 50	2,2,2,2		0.029	0.037	0.026	0.060	0.059	0.051	0.060
	5,5,5,5		0.029	0.037	0.026	0.060	0.059	0.051	0.060
30,40,60,70	2,2,2,2		0.041	0.055	0.038	0.058	0.057	0.047	0.056
	5,5,5,5		0.041	0.055	0.038	0.058	0.057	0.047	0.056
20,40,60,80	2,2,2,2		0.044	0.050	0.029	0.052	0.048	0.041	0.042
	5,5,5,5		0.044	0.050	0.029	0.052	0.048	0.041	0.042

Table 3. Type I error probability for k = 5

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
10,10,10,10,10	2,2,2,2,2	1, 1, 1, 1, 1	0.005	0.008	0.002	0.063	0.061	0.016	0.055
	5,5,5,5,5		0.005	0.008	0.002	0.063	0.061	0.016	0.055
6,8,10,12,14	2,2,2,2,2		0.006	0.019	0.001	0.070	0.073	0.020	0.049
	5,5,5,5,5		0.006	0.019	0.001	0.070	0.073	0.020	0.049
4,6,10,14,16	2,2,2,2,2		0.017	0.044	0.010	0.093	0.096	0.028	0.040
	5,5,5,5,5		0.017	0.044	0.010	0.093	0.096	0.028	0.040
30,30,30,30,30	2,2,2,2,2		0.022	0.036	0.012	0.052	0.053	0.037	0.052
	5,5,5,5,5		0.022	0.036	0.012	0.052	0.053	0.037	0.052
18,24,30,36,42	2,2,2,2,2		0.038	0.055	0.025	0.057	0.058	0.044	0.052
	5,5,5,5,5		0.038	0.055	0.025	0.057	0.058	0.044	0.052
12,18,30,42,48	2,2,2,2,2		0.040	0.047	0.025	0.050	0.051	0.034	0.033
	5,5,5,5,5		0.040	0.047	0.025	0.050	0.051	0.034	0.033
50,50,50, 50,50	2,2,2,2,2		0.030	0.043	0.024	0.046	0.047	0.038	0.047
	5,5,5,5,5		0.030	0.043	0.024	0.046	0.047	0.038	0.047
30,40,50,60,70	2,2,2,2,2		0.043	0.066	0.040	0.053	0.050	0.039	0.046
	5,5,5,5,5		0.043	0.066	0.040	0.053	0.050	0.039	0.046
20,30,50,70,80	2,2,2,2,2		0.042	0.046	0.024	0.043	0.042	0.037	0.036
	5,5,5,5,5		0.042	0.046	0.024	0.043	0.042	0.037	0.036

close to the nominal level. PB test in unbalanced designs appears to be less likely than the nominal level of Type I error probability in the relatively less unbalanced designs of medium and large samples and unbalanced designs of large samples. The increasing of the scale parameter a_i does not affect the controlling Type I error probability of the tests. The results of the Type I error probability of the tests k = 5 are given in Table 3.

The results given in Table 2 are examined, the increasing of the k decreased the Type 1 error probability of the tests. Type 1 error probability of CATs increased and it is more liberal than k = 4. Type 1 error probability of GP, PB, and FA tests decreases and gets closer to zero. Similar to the results in Table 3, GP, PB, FA, and CAT3 tests are conservative, while CAT1 and CAT2 tests are liberal in terms of Type I error probability. CAT4 and PB test's Type I error probability are closer than the others to the nominal level in small samples. It is observed that the Type I error probability of CAT1, CAT2, CAT4, and PB tests are close to the nominal level in medium samples. Type 1 error probability of PB, CAT1, CAT2, and CAT4 tests are close to the nominal level in large samples. The unbalancedness of the design does not affect the Type I error probability of the GP, PB, and FA test, while it affects CAT1, CAT2, and CAT4 negatively in small samples. In medium and large samples, the unbalancedness of the design only affects the ability to control the Type I error probability of CAT4. The increase of the scale parameter a_i does not affect the ability to control the Type 1 error probability.

According to the results in Tables 1, 2, and 3, it is observed that the ability to control the Type I error probability of the tests is affected by the sample sizes, number of populations and the design type. While the CAT1 and CAT2 control Type I error probability better k = 3, the situation of increasing in the k inflates the Type I error probability of these tests. When the increasing of the k, the Type I error probability of the CAT4 is closest to the nominal level. The unbalancedness of design does not affect the Type I error probability of the GP, PB, and FA test, while it affects the ability to control the Type I error probability of CATs negatively. It is observed that the Type I error probability of CATs may be sensitive to the unbalanced designs. Also, it is observed that the increase in the scale parameter does not affect the performance of the tests. According to the results obtained in this section, no test can ideally control Type 1 error probability in every case.

Table 4. The penalized power for lower scale parameters and k = 3

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
	2,2,3		0.0101	0.0143	0.0021	0.0847	0.0752	0.0226	0.0691
10,10,10	2,2,4		0.0254	0.0334	0.0055	0.0991	0.0919	0.0234	0.0808
	2,2,5		0.0429	0.0581	0.0097	0.1145	0.1012	0.0317	0.0898
	2,2,3		0.0245	0.0342	0.0039	0.0839	0.0773	0.0256	0.0635
8,10,12	2,2,4		0.0622	0.0779	0.0096	0.1077	0.1095	0.0380	0.0939
	2,2,5		0.1181	0.1386	0.0177	0.1135	0.1160	0.0334	0.0967
	2,2,3		0.0559	0.1119	0.0147	0.0579	0.0733	0.0230	0.0407
5,10,15	2,2,4		0.1590	0.2180	0.0265	0.0693	0.0875	0.0213	0.0371
	2,2,5		0.3010	0.3475	0.0413	0.0658	0.0808	0.0179	0.0281
	2,2,3		0.1610	0.2420	0.1253	0.3060	0.2901	0.2272	0.3021
30,30,30	2,2,4		0.4997	0.6363	0.4303	0.6721	0.6567	0.5812	0.6731
	2,2,5		0.7297	0.8488	0.6697	0.9004	0.8759	0.8103	0.8984
	2,2,3		0.2444	0.3225	0.1755	0.3040	0.3175	0.2734	0.3166
24,30,36	2,2,4	1, 1, 1	0.6301	0.7455	0.5232	0.7469	0.7418	0.6895	0.7457
	2,2,5		0.8083	0.9180	0.7411	0.9276	0.9153	0.8773	0.9228
	2,2,3		0.3307	0.4043	0.2160	0.3528	0.3737	0.3477	0.3381
15,30,45	2,2,4		0.7178	0.8277	0.5868	0.8019	0.8045	0.7852	0.7736
	2,2,5		0.8447	0.9643	0.7630	0.9296	0.9179	0.9232	0.9160
	2,2,3		0.3808	0.4686	0.3422	0.4672	0.4617	0.4604	0.4701
50,50,50	2,2,4		0.8119	0.9107	0.7675	0.8953	0.8795	0.9189	0.8963
	2,2,5		0.8812	0.9700	0.8447	0.9477	0.9310	0.9842	0.9477
	2,2,3		0.4722	0.5521	0.4015	0.5091	0.5283	0.5276	0.5332
40,50,60	2,2,4		0.8448	0.9522	0.8029	0.9096	0.9228	0.9335	0.9315
	2,2,5		0.8821	0.9921	0.8549	0.9506	0.9594	0.9776	0.9684
	2,2,3		0.5582	0.6071	0.4440	0.5556	0.5927	0.5562	0.5654
25,50,75	2,2,4		0.8840	0.9662	0.8103	0.9052	0.9257	0.8951	0.9144
	2,2,5		0.9063	0.9937	0.8489	0.9440	0.9623	0.9266	0.9525

Penalized Power Results

In addition to the configuration parameters used in Type 1 error probability calculations, the a_3 parameter is used to control the effect size in this section. For the low and high levels of the scale parameter, the penalized power of the test is calculated by fixing the values of the scale parameters; *3,4,5* and *6,8,10* respectively. In all scenarios, location parameters are taken as $b = 1,1,1$. The penalized powers are given for the low value of the scale parameter ($a_i = 2$) in Table 4, for the high value of the scale parameter ($a_i = 5$) in Table 5 for k = 3. The penalized powers are given for the low value of the scale parameter ($a_i = 2$) in Table 6, for the high value of the scale parameter ($a_i = 5$) in Table 7 for k = 4. The penalized powers are given for the low value of the scale parameter ($a_i = 2$) in Table 8, for the high value of the scale parameter ($a_i = 5$) in Table 9 for k = 5.

CAT1 is the most powerful test in balanced and unbalanced designs, the PB is the most powerful test in more unbalanced designs in terms of penalized power according to the results given in Tables 4 and 5. While the sample size increasing, the CAT4 is getting more powerful in a balanced and unbalanced design, the CAT2 is getting also more powerful than the others in a more unbalanced design. The increasing of the scale parameter value decreases the power

Table 5. The penalized power for higher scale parameters and k = 3

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
	5,5,6		0.0215	0.0234	0.0164	0.0599	0.0586	0.0118	0.0524
10,10,10	5,5,8		0.0298	0.0321	0.0266	0.0880	0.0850	0.0299	0.0810
	5,5,10		0.0419	0.0436	0.0369	0.0965	0.0935	0.0322	0.0858
	5,5,6		0.0356	0.0367	0.0349	0.0620	0.0644	0.0210	0.0525
8,10,12	5,5,8		0.0651	0.0642	0.0622	0.0858	0.0893	0.0303	0.0672
	5,5,10		0.1046	0.1051	0.1023	0.1077	0.1095	0.0380	0.0939
	5,5,6		0.0221	0.0631	0.0094	0.0623	0.0770	0.0247	0.0416
5,10,15	5,5,8		0.0726	0.1325	0.0171	0.0614	0.0910	0.0170	0.0398
	5,5,10		0.1577	0.2186	0.0266	0.0693	0.1080	0.0187	0.0444
	5,5,6		0.1509	0.1628	0.1329	0.0991	0.0954	0.0779	0.0991
30,30,30	5,5,8		0.4060	0.4592	0.3732	0.3769	0.3619	0.3042	0.3759
	5,5,10		0.5755	0.6553	0.5404	0.6721	0.6567	0.5812	0.6731
	5,5,6		0.0758	0.0868	0.0742	0.1166	0.1173	0.1040	0.1212
24,30,36	5,5,8	1, 1, 1	0.2967	0.3371	0.2831	0.4361	0.4443	0.4041	0.4436
	5,5,10		0.5172	0.5825	0.4921	0.7615	0.7656	0.7088	0.7736
	5,5,6		0.0886	0.1306	0.0588	0.1154	0.1236	0.1115	0.1058
15,30,45	5,5,8		0.4225	0.4979	0.2934	0.4605	0.4720	0.4365	0.4327
	5,5,10		0.7141	0.8159	0.5891	0.8019	0.8045	0.7852	0.7736
	5,5,6		0.2984	0.3276	0.2862	0.1211	0.1199	0.1059	0.1211
50,50,50	5,5,8		0.6130	0.6734	0.5938	0.5988	0.5872	0.5921	0.5997
	5,5,10		0.6634	0.7277	0.6399	0.8953	0.8795	0.9189	0.8963
	5,5,6		0.1163	0.1323	0.1123	0.1411	0.1511	0.1432	0.1505
40,50,60	5,5,8		0.4950	0.5549	0.4724	0.6407	0.6591	0.6599	0.6712
	5,5,10		0.6630	0.7370	0.6277	0.9096	0.9228	0.9335	0.9315
	5,5,6		0.1498	0.1752	0.1097	0.1713	0.1814	0.1611	0.1602
25,50,75	5,5,8		0.6815	0.7300	0.5755	0.7099	0.7266	0.6785	0.6674
	5,5,10		0.8951	0.9605	0.8291	0.9486	0.9433	0.8874	0.8828

of the PB test. In the results given in Tables 6 and 7, it is observed that the CAT1, CAT2, and CAT4 are most powerful in balanced and unbalanced designs, the GP and PB test have similar power in unbalanced designs. A significant increase in the power of CAT3 is observed and it is one of the most powerful tests when the sample size increases. However, the increasing of scale parameter value decreases the power of the GP test, the PB, CAT3, and CAT4 maintain their performance. Results k = 5 given in Tables 8 and 9 indicate that the CAT1 and CAT3 are the most powerful, also the GP and PB tests are most powerful in only unbalanced designs. When the sample size increases and the design type is unbalanced, the performance of the CAT1 and CAT3 decreases, and the GPV and PB tests maintain their powers. Despite the increasing scale parameter value, the GP, PB, and CAT3 maintain their powers. As a result, CATs outperform others in balanced design, the PB test is the most powerful in unbalanced design. When the k increases, the power of the GP test is competitive with the powerful tests. The PB and CAT1 tests maintain their power against the increase in the value of the scale parameter.

Table 6. The penalized power for lower scale parameters and $k = 4$

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
	2,2,2,3		0.0096	0.0135	0.0028	0.0864	0.0847	0.0293	0.0780
10,10,10,10	2,2,2,4		0.0170	0.0315	0.0036	0.0990	0.0942	0.0372	0.0928
	2,2,2,5		0.0229	0.0457	0.0057	0.0982	0.0933	0.0356	0.0854
	2,2,2,3		0.0243	0.0375	0.0043	0.0757	0.0859	0.0250	0.0726
6,8,12,14	2,2,2,4		0.0678	0.0804	0.0072	0.0765	0.0835	0.0306	0.0775
	2,2,2,5		0.1445	0.1249	0.0079	0.0765	0.0788	0.0290	0.0618
	2,2,2,3		0.0474	0.0713	0.0058	0.0678	0.0757	0.0263	0.0418
4,8,12,16	2,2,2,4		0.1468	0.1220	0.0094	0.0671	0.0772	0.0214	0.0409
	2,2,2,5		0.2861	0.1978	0.0152	0.0580	0.0645	0.0214	0.0334
	2,2,2,3		0.1296	0.2041	0.0919	0.1961	0.2013	0.1534	0.2023
30,30,30,30	2,2,2,4		0.4103	0.5641	0.3435	0.5505	0.5679	0.4546	0.5629
	2,2,2,5		0.7006	0.8315	0.6133	0.8046	0.8259	0.7175	0.8204
	2,2,2,3		0.2992	0.3359	0.1655	0.3110	0.3275	0.2626	0.2910
18,24,36,42	2,2,2,4	1, 1, 1, 1	0.7183	0.7578	0.5341	0.7800	0.7953	0.6917	0.7427
	2,2,2,5		0.8200	0.8845	0.7325	0.9620	0.9512	0.8673	0.9118
	2,2,2,3		0.3414	0.3565	0.1578	0.3349	0.3465	0.3069	0.2990
12,24,36,48	2,2,2,4		0.7551	0.8169	0.5513	0.7727	0.7783	0.7703	0.7693
	2,2,2,5		0.8385	0.9723	0.7846	0.9372	0.9141	0.9565	0.9614
	2,2,2,3		0.3063	0.3813	0.2713	0.3670	0.3719	0.3693	0.3679
50,50,50, 50	2,2,2,4		0.7276	0.8080	0.6941	0.8271	0.8350	0.8832	0.8289
	2,2,2,5		0.8350	0.8891	0.8154	0.9083	0.9160	0.9842	0.9074
	2,2,2,3		0.5616	0.5568	0.4221	0.5153	0.5376	0.5488	0.5367
30,40,60,70	2,2,2,4		0.9058	0.9325	0.8639	0.9062	0.9179	0.9509	0.9251
	2,2,2,5		0.9206	0.9525	0.8971	0.9285	0.9366	0.9713	0.9449
	2,2,2,3		0.6425	0.6050	0.4171	0.5834	0.5982	0.5422	0.5376
20,40,60,80	2,2,2,4		0.9317	0.9720	0.8048	0.9600	0.9629	0.8994	0.9071
	2,2,2,5		0.9449	1	0.8383	0.9806	0.9806	0.9206	0.9285

REAL DATA APPLICATION

In this application, the effect of antidotes developed against the poisons was tested on mice. The data set poisons obtained from [27], and available in boot package in R, the response times of the mice against the treatment of antidotes A, B, C, and D were made. Parameter estimates and summary statistics (in week) are given in Table 10.

According to the summary statistics given in Table 10, antidote treatment is given to mice. The mean duration of remission is *0.33* weeks in antidote A, *0.71* in antidote B, and antidote C hence *0.41* weeks in antidote and *0.56* weeks in antidote. It is observed that the estimates of the scale parameters are different. The data set is an example of a small sample balanced design. The distribution of remission times to treatment is given in Figure 1.

The distribution of the remission times of the poisons are right skewed and not normal according to the p-value of the Shapiro-Wilk Normality test is *0.0001* at the significance level is *0.05*. Thus, the proposed tests for two-parameter exponential distribution should be used because of the distribution of data. In this example, the aim is to test whether antidotes have equal effects on the duration of remission. Thus, the most effective treatment can be

Table 7. The penalized power for higher scale parameters and $k = 4$

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
	5,5,5,6		0.0066	0.0090	0.0021	0.0758	0.0709	0.0241	0.0703
10,10,10,10	5,5,5,8		0.0125	0.0172	0.0028	0.0977	0.0893	0.0342	0.0901
	5,5,5,10		0.0170	0.0315	0.0036	0.1077	0.1003	0.0365	0.0990
	5,5,5,6		0.0118	0.0187	0.0029	0.0684	0.0748	0.0236	0.0566
6,8,12,14	5,5,5,8		0.0317	0.0492	0.0043	0.0780	0.0929	0.0244	0.0687
	5,5,5,10		0.0678	0.0804	0.0072	0.0780	0.0863	0.0299	0.0733
	5,5,5,6		0.0163	0.0463	0.0044	0.0664	0.0716	0.0279	0.0383
4,8,12,16	5,5,5,8		0.0667	0.0820	0.0065	0.0640	0.0753	0.0270	0.0429
	5,5,5,10		0.1468	0.1220	0.0094	0.0679	0.0768	0.0270	0.0402
	5,5,5,6		0.0397	0.0737	0.0281	0.0702	0.0718	0.0543	0.0728
30,30,30,30	5,5,5,8		0.1745	0.2684	0.1299	0.2800	0.2873	0.2090	0.2816
	5,5,5,10		0.4103	0.5641	0.3435	0.5601	0.5679	0.4584	0.5679
	5,5,5,6		0.0683	0.0960	0.0351	0.1089	0.1127	0.0886	0.1002
18,24,36,42	5,5,5,8	1, 1, 1, 1	0.3983	0.4463	0.2280	0.3951	0.4215	0.3394	0.3846
	5,5,5,10		0.7183	0.7578	0.5341	0.7723	0.7877	0.6701	0.7427
	5,5,5,6		0.0783	0.1109	0.0453	0.1049	0.1200	0.0945	0.0942
12,24,36,48	5,5,5,8		0.4444	0.4624	0.2325	0.4233	0.4384	0.3846	0.3973
	5,5,5,10		0.7551	0.8169	0.5513	0.7521	0.7852	0.7351	0.7547
	5,5,5,6		0.0621	0.0873	0.0526	0.1178	0.1167	0.0966	0.1158
50,50,50, 50	5,5,5,8		0.4137	0.5060	0.3748	0.5347	0.5276	0.4707	0.5287
	5,5,5,10		0.7276	0.8080	0.6971	0.8971	0.8894	0.8282	0.8991
	5,5,5,6		0.1464	0.1621	0.1069	0.1451	0.1545	0.1346	0.1424
30,40,60,70	5,5,5,8		0.6867	0.6817	0.5649	0.6776	0.7040	0.6406	0.6678
	5,5,5,10		0.9058	0.9325	0.8639	0.9570	0.9703	0.9090	0.9420
	5,5,5,6		0.1446	0.1570	0.0831	0.1602	0.1575	0.1545	0.1471
20,40,60,80	5,5,5,8		0.7644	0.7310	0.5379	0.7046	0.6790	0.7278	0.7148
	5,5,5,10		0.9317	0.9720	0.8048	0.9334	0.8891	0.9674	0.9580

provided by determining the antidotes or antidotes to be used. The tests are considered in this study implemented in doex package as well as the other ANOVA methods under non-normality [28, 29, 30] and used for testing null hypothesis and the p-values obtained are given in Table 11.

The null hypothesis is rejected at the *0.05* significance level according to the p-value results obtained from all tests except the FA test from the tests given in the table above. Thus, it is concluded that the duration of remission of antidotes is not equal.

CONCLUSION

Based on the simulation results and the application to a real dataset, several conclusions can be drawn regarding the performance of the proposed tests and their applicability in practical scenarios. In both balanced and unbalanced designs, CAT1 emerges as the most powerful test, outperforming others. The PB test, while generally powerful, exhibits its highest power in more unbalanced designs, showcasing its suitability for specific scenarios. As the sample size increases, CAT4 becomes more powerful in both balanced and unbalanced designs, with CAT2

Table 8. The penalized power for lower scale parameters and k = 5

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
	2,2,2,2,3		0.0065	0.0170	0.0014	0.0815	0.0170	0.0744	0.0576
10,10,10,10,10	2,2,2,2,4		0.0123	0.0280	0.0029	0.0969	0.0216	0.0849	0.0685
	2,2,2,2,5		0.0210	0.0383	0.0036	0.1041	0.0208	0.0791	0.0736
	2,2,2,2,3		0.0204	0.0322	0.0064	0.0753	0.0213	0.0624	0.0533
6,8,10,12,14	2,2,2,2,4		0.0430	0.0605	0.0078	0.0844	0.0190	0.0713	0.0597
	2,2,2,2,5		0.0955	0.0951	0.0092	0.0753	0.0134	0.0525	0.0533
	2,2,2,2,3		0.0458	0.0756	0.0119	0.0779	0.0283	0.0438	0.0551
4,6,10,14,16	2,2,2,2,4		0.1288	0.1266	0.0157	0.0765	0.0250	0.0438	0.0541
	2,2,2,2,5		0.2616	0.1587	0.0171	0.0628	0.0200	0.0338	0.0444
	2,2,2,2,3		0.0905	0.1564	0.0618	0.1991	0.1363	0.2049	0.1408
30,30,30,30,30	2,2,2,2,4		0.3139	0.4552	0.2329	0.5206	0.3751	0.5226	0.3681
	2,2,2,2,5		0.5741	0.7239	0.4741	0.7557	0.5960	0.7698	0.5343
	2,2,2,2,3		0.3062	0.3061	0.1511	0.2785	0.2504	0.2726	0.1970
18,24,30,36,42	2,2,2,2,4	1, 1, 1, 1, 1	0.7292	0.7323	0.4736	0.7112	0.6699	0.7325	0.5029
	2,2,2,2,5		0.8738	0.9153	0.7201	0.8913	0.8825	0.9316	0.6303
	2,2,2,2,3		0.3761	0.3186	0.1560	0.3228	0.2193	0.2272	0.2282
12,18,30,42,48	2,2,2,2,4		0.8079	0.7848	0.4793	0.7713	0.6040	0.6203	0.5454
	2,2,2,2,5		0.9074	0.9480	0.7422	0.9545	0.8121	0.8146	0.6749
	2,2,2,2,3		0.2688	0.3578	0.2239	0.3885	0.3269	0.3875	0.2747
50,50,50, 50,50	2,2,2,2,4		0.7091	0.8308	0.6578	0.8635	0.7705	0.8654	0.6106
	2,2,2,2,5		0.8384	0.9300	0.8038	0.9625	0.8881	0.9635	0.6806
	2,2,2,2,3		0.5460	0.4630	0.3871	0.5510	0.4817	0.5138	0.3896
30,40,50,60,70	2,2,2,2,4		0.9160	0.8408	0.8544	0.9730	0.8800	0.9353	0.6880
	2,2,2,2,5		0.9356	0.8686	0.9092	0.9980	0.9035	0.9603	0.7057
	2,2,2,2,3		0.6406	0.5485	0.3739	0.5283	0.4802	0.4702	0.3726
20,30,50,70,80	2,2,2,2,4		0.9210	0.9372	0.7706	0.9081	0.8695	0.8636	0.6421
	2,2,2,2,5		0.9285	0.9613	0.8103	0.9266	0.8891	0.8821	0.6552

also demonstrating increased power, particularly in more unbalanced designs. However, it's noteworthy that the power of the PB test decreases with an increase in the scale parameter value.

Considering the results for various scenarios and design types, CAT1, CAT2, and CAT4 consistently prove to be the most powerful tests, while GP and PB tests demonstrate similar power in unbalanced designs. CAT3 exhibits a significant increase in power with an increase in the sample size, becoming one of the most powerful tests. Notably, GP's power decreases with an increase in the scale parameter value, whereas PB, CAT3, and \textbf{CAT4} maintain their performance.

For the case when k = 5, CAT1 and CAT3 emerge as the most powerful tests, with GP and PB being particularly powerful in unbalanced designs. With an increase in sample size and unbalanced design, the power of CAT1 and CAT3 decreases, but GP and PB tests maintain their effectiveness. Remarkably, despite an increase in the scale parameter value, GP, PB, and CAT3 maintain their power. CATs consistently outperform other tests in balanced designs, while the PB test stands out as the most powerful test in unbalanced designs. As k increases, GP's power competes

Table 9. The penalized power for higher scale parameters and $k = 5$

n_i	a_i	b_i	GP	PB	FA	CAT1	CAT2	CAT3	CAT4
	5,5,5,5,6		0.0036	0.0088	0.0014	0.0625	0.0147	0.0553	0.0442
10,10,10,10,10	5,5,5,5,8		0.0080	0.0177	0.0021	0.0869	0.0193	0.0839	0.0615
	5,5,5,5,10		0.0123	0.0280	0.0029	0.0969	0.0216	0.0849	0.0685
	5,5,5,5,6		0.0066	0.0212	0.0036	0.0662	0.0182	0.0574	0.0468
6,8,10,12,14	5,5,5,5,8		0.0226	0.0401	0.0064	0.0778	0.0213	0.0634	0.0550
	5,5,5,5,10		0.0430	0.0605	0.0078	0.0844	0.0190	0.0713	0.0597
	5,5,5,5,6		0.0248	0.0539	0.0089	0.0736	0.0300	0.0402	0.0521
4,6,10,14,16	5,5,5,5,8		0.0636	0.0841	0.0127	0.0808	0.0300	0.0466	0.0572
	5,5,5,5,10		0.1288	0.1266	0.0157	0.0765	0.0250	0.0438	0.0541
	5,5,5,5,6		0.0280	0.0530	0.0204	0.0728	0.0561	0.0765	0.0515
30,30,30,30,30	5,5,5,5,8		0.1201	0.2024	0.0087	0.2574	0.1782	0.2657	0.1820
	5,5,5,5,10		0.3139	0.4552	0.2329	0.5206	0.3751	0.5226	0.3681
	5,5,5,5,6		0.0727	0.1106	0.0425	0.0975	0.0841	0.0951	0.0689
18,24,30,36,42	5,5,5,5,8	1, 1, 1, 1, 1	0.4050	0.4014	0.2107	0.3630	0.3184	0.3716	0.2567
	5,5,5,5,10		0.7292	0.7323	0.4736	0.7112	0.6699	0.7325	0.5029
	5,5,5,5,6		0.1050	0.1166	0.0449	0.1059	0.0679	0.0708	0.0749
12,18,30,42,48	5,5,5,5,8		0.4957	0.4089	0.2009	0.4198	0.2916	0.3032	0.2969
	5,5,5,5,10		0.8079	0.7848	0.4793	0.7713	0.6040	0.6203	0.5454
	5,5,5,5,6		0.0625	0.0927	0.0527	0.1020	0.0826	0.1010	0.0721
50,50,50, 50,50	5,5,5,5,8		0.3651	0.4730	0.3188	0.5012	0.4328	0.5090	0.3544
	5,5,5,5,10		0.7091	0.8308	0.6578	0.8635	0.7705	0.8654	0.6106
	5,5,5,5,6		0.1264	0.1271	0.0895	0.1290	0.1059	0.1174	0.0912
30,40,50,60,70	5,5,5,5,8		0.6715	0.5762	0.5121	0.6860	0.6066	0.6563	0.4851
	5,5,5,5,10		0.9160	0.8408	0.8544	0.9730	0.8800	0.9353	0.6880
	5,5,5,5,6		0.1569	0.1453	0.0819	0.1337	0.1167	0.1131	0.0945
20,30,50,70,80	5,5,5,5,8		0.7614	0.6909	0.4826	0.6629	0.6076	0.6028	0.4688
	5,5,5,5,10		0.9210	0.9372	0.7706	0.9081	0.8695	0.8636	0.6421

Table 10. The summary statistics of the data

Statistics	A	B	C	D
n_i	12	12	12	12
$\hat{a}_i$	0.15	0.42	0.19	0.26
$\hat{b}_i$	0.18	0.29	0.22	0.30
$\hat{\mu}_i$	0.33	0.71	0.41	0.56

Table 11. The results of the tests

Test	p-value
GP	0.0475*
PB	0.0205*
FA	0.0630
CAT1	0.0032**
CAT2	0.0038**
CAT3	0.0077**
CAT4	0.0036**

*significance at $\alpha = 0.05$, **significance at $\alpha = 0.01$

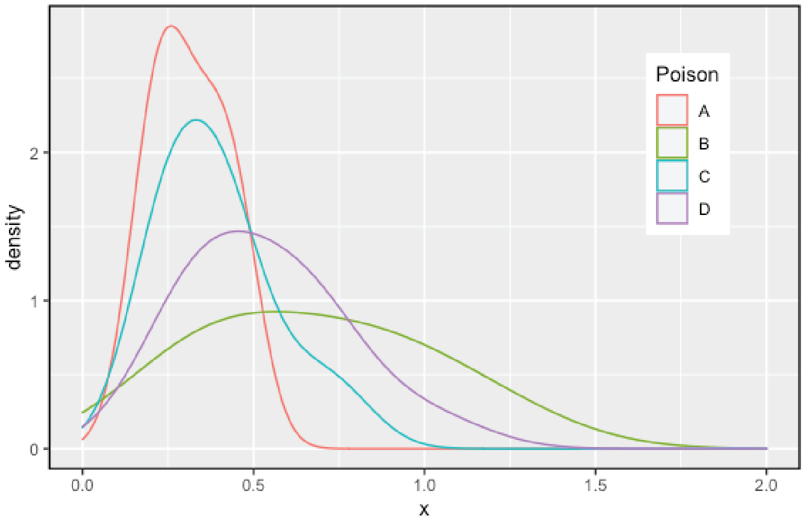

Figure 1. The PDFs of remission times of the poisons.

with other powerful tests. Both PB and CAT1 tests maintain their power against an increase in the scale parameter value. In conclusion, the proposed computational approach tests, particularly CAT1 and PB, demonstrate robust performance in various simulated scenarios and a real-world application involving antidote treatment on mice. These findings highlight the efficacy of the proposed tests in handling skewed distributions, emphasizing their potential in addressing Behrens-Fisher-type problems in non-normal settings.

ACKNOWLEDGEMENTS

This study is supported by the Eskisehir Technical University Scientific Research Projects Commission under grant No. 20DRP047.

AUTHORSHIP CONTRIBUTIONS

Authors equally contributed to this work.

DATA AVAILABILITY STATEMENT

The authors confirm that the data that supports the findings of this study are available within the article. Raw data that support the finding of this study are available from the corresponding author, upon reasonable request.

CONFLICT OF INTEREST

The author declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

ETHICS

There are no ethical issues with the publication of this manuscript.

REFERENCES

[1] Weerahandi S. ANOVA under unequal error variances. Biometrics 1995;51:589–599. [CrossRef]

[2] Krishnamoorthy K, Lu F, Mathew T. A parametric bootstrap approach for ANOVA with unequal variances: Fixed and random models. Comput Stat Data Anal 2007;51:5731–5742. [CrossRef]

[3] Li X, Wang J, Liang H. Comparison of several means: A fiducial based approach. Comput Stat Data Anal 2011;55:1993–2002. [CrossRef]

[4] Gokpinar EY, Gokpinar F. A test based on the computational approach for equality of means under the unequal variance assumption. Hacettepe J Math Stat 2012;41:605–613.

[5] Gokpinar EY, Gokpinar F. Testing the equality of several log-normal means based on a computational approach. Commun Stat Simul Comput 2017;46:1998–2010. [CrossRef]

[6] Jafari AA, Abdollahnezhad K. Testing the equality means of several lognormal distributions, Commun Stat Simul Comput 2017;46:2311–2320. [CrossRef]

[7] Weerahandi S, Krishnamoorthy K. A note reconciling ANOVA tests under unequal error variances. Comput Stat Data Anal 2019;55:1993–2002. [CrossRef]

[8] Chen HJ. A new range statistic for comparisons of several exponential location parameters. Biometrika 1982;69:257–260. [CrossRef]

[9] Singh N. The likelihood ratio test for the equality of location parameters of exponential populations

based on Type II censored samples. Technometrics 1983;25:193–195. [CrossRef]

[10] Kambo NS, Awad AM. Testing equality of location parameters of k exponential distributions. Commun Stat Theory Methods 1985;14:567–583.

[11] Hsieh HK. An exact test for comparing location parameters of k exponential distributions with unequal scaled based on Type II censored data. Technometrics 1986;28:157–164. [CrossRef]

[12] Vaughan DC, Tiku ML. Testing the equality of location parameters of exponential populations from censored samples. Commun Stat Theory Methods 1993;22:2567–2581. [CrossRef]

[13] Tiku ML, Vaughan DC. Testing equality of location parameters of two exponential distributions from censored samples. Commun Stat Theory Methods 1991;20:929–944. [CrossRef]

[14] Ananda MMA, Weerahandi S. Testing the difference of two exponential means using generalized p-values. Commun Stat Simul Comput 1996;25:521–532. [CrossRef]

[15] Wu SF. One stage multiple comparisons with the control for exponential mean lifetimes based on doubly censored samples under heteroscedasticity. Commun Stat Simul Comput 2021;50:1473–1483. [CrossRef]

[16] Malekzadeh A, Jafari AA. Inference on the equality means of several two-parameter exponential distributions under progressively Type II censoring, Commun Stat Simul Comput 2020;49:3196–3211. [CrossRef]

[17] Johnson NL, Kotz S. Continuous Univarite Distributions. Boston: Houghton Miffin Company; 1970.

[18] Raiz M, Kumar A, Mishra VN, Rao N. Dunkl analogue of Schurer-Beta operators and their approximation behaviour. Math Found Comput 2022;5:315–330. [CrossRef]

[19] Rajawat RS, Singh KK, Mishra VN. Approximation by modified Bernstein polynomials based on real parameters. Math Found Comput 2024;7:297–309. [CrossRef]

[20] Cochran WG. Problem arising in the analysis of a series of similar experiments. J Royal Stat Soc 1937;4:102–118. [CrossRef]

[21] Tsui KW, Weerahandi S. Generalized p-value in significance testing of hypothesis in the presence of nuisance parameters. J Am Stat Assoc 1989;84:602–607. [CrossRef]

[22] Pal N, Lim WK, Ling CH. A computational approach to statistical inferences. J Appl Probab Stat 2007;2:13–35.

[23] Zheng M. Penalized maximum likelihood estimation of two-parameter exponential distributions. Minnesota, USA: The Faculty of Graduate School of the University of Minnesota; 2013.

[24] Cavus M, Yazici B, Sezer A. Penalized power approach to compare the power of the tests when Type I error probabilities are different. Commun Stat Simul Comput 2021;50:1912–1926. [CrossRef]

[25] Cavus M, Yazici B. Doex: One-way heteroscedastic ANOVA tests. R package: ver.1.3, 2020.

[26] Cavus M, Yazici B. Testing the equality of normal distributed and independent groups' means under unequal variances by doex package. The R J 2021;12:134–154. [CrossRef]

[27] Box GEP, Cox DR. An analysis of transformation. J Royal Stat Soc 1964;26:211–252. [CrossRef]

[28] Cavus M, Yazici B, Sezer A. Modified tests for comparison of group means under heteroskedasticity and non-normality caused by outlier(s). Hacettepe J Math Stat 2017;46:493–510. [CrossRef]

[29] Cavus M, Yazici B, Sezer A. Analyzing regional export data by the modified generalized F-test. Int J Econ Administr Stud 2018:541–552.

[30] Cavus M, Yazici B, Sezer A. Penalized power approach to compare the power of the tests when Type I error probabilities are different. Commun Stat Simul Comput 2021;50:1912–1926. [CrossRef]