

*Zofia Hanusz**

RELATIVE POTENCY FOR THE MULTIVARIATE CONTAMINATED NORMAL RESPONSES

Abstract. In this paper we focus on the impact of responses of contaminated normal distribution on the relative potency. For several values of the contamination parameters, the estimates of the relative potency, its goodness and the truthfulness of the hypotheses connected with the estimation of the relative potency are tabulated for the generated data sets.

1. INTRODUCTION

In biological assays, comparing two preparations: standard (known) and test (new) we get the multivariate responses. In this case the relative potency is frequently estimated. In the multivariate setting, to derive the estimator of the relative potency we have to assume that the responses are normally distributed. In practice, however, this assumption is not necessarily fulfilled. It is of interest to study how the estimates of the relative potency differ from the true value of the parameter in the cases where responses do not fulfil the normality assumption. In this paper, we concentrate only on the contaminated multivariate normal distribution of the observations which is more frequently encountered. On the generated data sets, for several values of the contamination parameters, the average estimates of the relative potency, standard deviations of these estimates, probability of acceptance and the hypotheses connected with the estimation are tabulated.

* Agricultural University of Lublin, Institute of Applied Mathematics.

2. STATEMENT OF THE PROBLEM

In this section we recall the problem of estimation of the relative potency in the multivariate setting. Let us consider an experiment with two preparations: Standard denoted by S and Test denoted by T . Let the preparations be applied on v_i ($i = S, T$) doses: $u_{i1}, u_{i2}, \dots, u_{iv_i}$, which are repeated $n_{i1}, n_{i2}, \dots, n_{iv_i}$ times, respectively. For each dose of the preparations, the same p -features are measured as the p variate response. Without loss of generality, we can assume that the doses are applied to homogeneous units. Then, it is well known (see Finney, 1978) that each response y_{ijk} can be described as

$$y_{ijk} = \alpha_i + \beta_i x_{ij} + e_{ijk}, \quad i = S, T; \quad j = 1, \dots, v_i; \quad k = 1, \dots, n_{ij}$$

where α_i, β_i are $(p \times 1)$ vectors of intercepts and slopes, x_{ij} is the logarithm to base 10 of the dose u_{ij} . Note that the relative potency, denoted by ρ , is defined as the ratio of the dose of the Standard preparations to the dose of the Test preparations which give the same multivariate responses, so $\rho = \frac{u_T}{u_S}$. As y_{ijk} depends on the logarithms of the doses so we also consider the logarithm of the relative potency, denoted by μ , which is the distance between the logarithms of the proper doses of the preparations. The total model of the experiment thus defined can be written as follows:

$$\mathbf{Y} = \mathbf{XB} + \mathbf{E} \quad (1)$$

where $\mathbf{Y} = \begin{pmatrix} \mathbf{Y}_S \\ \mathbf{Y}_T \end{pmatrix}$, and \mathbf{Y}_i an $(n_i \times p)$ matrix, whose rows are equal to y'_{ijk} , $n_i = \sum_{j=1}^{v_i} n_{ij}$, $\mathbf{X} = \begin{pmatrix} \mathbf{1}_{n_S} & \mathbf{x}_S & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1}_{n_T} & \mathbf{x}_T \end{pmatrix}$, \mathbf{x}_i is an $(n_i \times 1)$ vector composed of all x_{ij} in the same order as the observations in the matrix \mathbf{Y}_i , $\mathbf{1}_{n_i}$ is the $(n_i \times 1)$ vector of ones, $\mathbf{B} = (\alpha_S, \beta_S, \alpha_T, \beta_T)'$, $\mathbf{E} = \begin{pmatrix} \mathbf{E}_S \\ \mathbf{E}_T \end{pmatrix}$ and \mathbf{E}_i is an $(n_i \times p)$ matrix of errors whose rows are mutually independent and normally distributed with null vector of expectations and the same covariance matrix Σ . Before we get on to the presentation of the results of estimation, we will briefly describe two fundamental hypotheses connected with the estimation of the hypothesis about the parallel-line designs and the hypothesis about the relative potency.

2.1. Hypothesis about a parallel-line design

The relative potency of two preparations is derived in so the called parallel-line designs, having the same vectors of slopes: β_S and β_T . The equality of the slopes is expressed as the following hypothesis:

$$H_\beta^0: C'B = 0' \quad \text{versus} \quad H_\beta^1: C'B \neq 0' \quad (2)$$

where $C' = (0, 1, 0, -1)$. To test (2) one can use *Wilks' lambda* statistic which is an F Snedecor statistic taking the form (see H a n u s z, 1998):

$$F^0 = \frac{n_S + n_T - r(X) - p + 1}{p} \cdot \frac{(C'\hat{B})S_E^{-1}(C'\hat{B})'}{C'(X'X)^{-1}C}$$

where $\hat{B} = (X'X)^{-1}X'Y$, $S_E = (Y - X\hat{B})(Y - X\hat{B})'$, $r(X)$ is the rank of X . The hypothesis H_β^0 in (2) has to be accepted to have parallel line design.

2.2. Hypothesis about the relative potency

Under the truthfulness of the null hypothesis in (2), the model (1) is reparametrized to the following model:

$$X = XB + E \quad (3)$$

where $X = \begin{pmatrix} \mathbf{1}_{n_S} & \mathbf{0} & \mathbf{x}_S \\ \mathbf{0} & \mathbf{1}_{n_T} & \mathbf{x}_T \end{pmatrix}$, $B = (\alpha_S, \alpha_T, \beta)'$; and α_S, α_T remain the same as in the model (1) but β is the common vector of the slopes. The main hypothesis about the logarithm of the relative potency μ is written in the form:

$$H_\mu^0: C'_\mu B = 0' \quad \text{versus} \quad H_\mu^1: C'_\mu B \neq 0' \quad (4)$$

where $C'_\mu = (1, -1, \mu)$. The hypothesis H_μ^0 is tested by *Wilks' lambda* statistic taking the form (see, H a n u s z, 1995):

$$\Lambda_\mu = \frac{1}{1 + V_\mu} \quad (5)$$

where

$$V_\mu = \frac{(\mathbf{C}'_\mu \hat{\mathbf{B}}) \mathbf{S}_E^{-1} (\mathbf{C}'_\mu \hat{\mathbf{B}})'}{\mathbf{C}'_\mu (\mathbf{X}'\mathbf{X})^{-1} \mathbf{C}_\mu},$$

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y},$$

$$\mathbf{S}_E = (\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}})'(\mathbf{Y} - \mathbf{X}\hat{\mathbf{B}}).$$

Considering the fact that $\left(n_S + n_T - r(\mathbf{X}) - \frac{p-1}{2} + \frac{1}{\min(\Lambda_\mu)} \right) \ln(1 + V_\mu)$ has approximately χ^2 distribution with $(p-1)$ degrees of freedom (see Williams, 1988), we can test the null hypothesis in (4). The estimator of the logarithm of the relative potency is such $\hat{\mu}$ which maximizes Λ_μ under the truthfulness of the of the null hypothesis in (4).

3. RESULTS FOR GENERATED DATA SETS FROM NORMAL DISTRIBUTION

Testing the hypotheses and estimation of the relative potency presented in the previous section are carried out assuming that all observations have a p -variate normal distribution. In this section we illustrate the results obtained for the generated data sets, having the normal distribution. In order to get the estimates of the logarithm of the relative potency and to observe the probabilities of the truthfulness of the null hypothesis in (2) and (4), the following parameters in the model (1) were fixed:

$$p = 3, \quad v_S = v_T = 3, \quad u_{ST} = u_{T1} = 250, \quad u_{S2} = u_{T2} = 500, \quad u_{S3} = u_{T3} = 1000,$$

$$\alpha_S = \begin{pmatrix} 1 \\ 22 \\ 100 \end{pmatrix}, \quad \alpha_T = \begin{pmatrix} 1,1 \\ 23.0 \\ 105.0 \end{pmatrix}, \quad \beta_S = \beta_T = \begin{pmatrix} 0.1 \\ 1.0 \\ 5.0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} 10 & -4 & 1 \\ -4 & 8 & 0 \\ 1 & 0 & 3 \end{pmatrix}.$$

Let us notice that the values of the model parameters are chosen in such a way that the logarithm of the relative potency μ is equal to one. Namely, the difference $\alpha_T - \alpha_S$ is the same as $\beta (= \beta_S = \beta_T)$, so the parameter μ of the null hypothesis in (4) has to be one. For the different numbers of dose replications, calculations were repeated 100 times for the generated data sets using MapleV package. The hypotheses were tested on 5 percent of the significant level. The results obtained by simulations are presented in Tab. 1.

Table 1

Probabilities of truthfulness hypotheses, average estimates of logarithm of relative potency and standard deviation of estimates for the normally distributed data

Dose replications	i) H_β^0	ii) H_μ^0	iii) $ \mu - \hat{\mu} < 0.2$	$i \wedge ii \wedge iii$	$\hat{\mu}$	S_μ
$n_S = 5, n_T = 5$	0.90	0.94	0.44	0.37	1.07	0.41
$n_S = 5, n_T = 10$	0.90	0.94	0.56	0.48	1.05	0.34
$n_S = 10, n_T = 10$	0.89	0.97	0.67	0.58	0.99	0.26
$n_S = 20, n_T = 20$	0.92	0.91	0.85	0.72	1.02	0.14
$n_S = 25, n_T = 25$	0.95	0.94	0.88	0.78	1.02	0.13
$n_S = 30, n_T = 30$	0.95	0.97	0.90	0.82	1.00	0.11

The second and the third columns of this Table enclose the probabilities of the truthfulness of the null hypotheses in (2) and (4). The fourth column contains the probability that the estimates of the logarithm of the relative potency differ from the true value of the logarithm of the relative potency less than by 0.2. The next column gives us the joint probability of the conditions of the three previous columns. In the last two columns we can see the average estimate of the logarithm of the relative potency and standard deviation of all estimates obtained through the simulations. Table 1 shows that the estimates are closer to the parameter when the number of dose replications is bigger. Regarding the probabilities of acceptance of the null hypotheses about parallelism and the logarithm of the relative potency we can notice that they are high enough even for the small number of dose replications. Looking at the last two columns, it is easy to conclude that the estimates of the logarithm of the relative potency are better when the number of dose replications is higher.

4. RESULTS FOR THE CONTAMINATED NORMAL DISTRIBUTED DATA

As we mentioned, in practice, however, the assumption about the normality of responses is not necessarily fulfilled. It is worth checking out whether the estimate of the relative potency differs from the true parameter when some of the responses have another distribution. In this paper we restrict our attention to the contaminated normal distribution. We concentrate on the situation where most of the responses have the distribution described in the model (1) but q percent of the data, chosen randomly, has the normal distribution with intercepts shifted by the vector \mathbf{a} and the covariance matrix $r\Sigma$. Results for the different q and r are enclosed in Tab. 2, 3, 4 and 5. The columns of the tables are constructed as in Tab. 1.

Table 2

The estimates of the logarithm of the relative potency and probabilities of testing hypotheses with $r = 0.5$

q	n_S	n_T	i) H_β^0	ii) H_μ^0	iii) $ \mu - \hat{\mu} < 0.2$	$i \wedge ii \wedge iii$	$\hat{\mu}$	$S_{\hat{\mu}}$
0.1	5	5	0.93	0.97	0.40	0.35	1.38	2.20
	5	10	0.88	0.94	0.44	0.36	1.18	0.53
	10	10	0.94	0.91	0.54	0.49	1.09	0.38
	20	20	0.93	0.96	0.71	0.63	1.03	0.22
	25	25	0.92	0.95	0.77	0.68	1.04	0.20
	30	30	0.92	0.93	0.81	0.72	1.00	0.15
0.2	5	5	0.92	0.93	0.33	0.30	1.55	2.79
	5	10	0.92	0.90	0.39	0.32	1.14	0.39
	10	10	0.90	0.96	0.39	0.32	1.18	0.47
	20	20	0.97	0.93	0.69	0.62	1.05	0.23
	25	25	0.93	0.91	0.71	0.61	1.01	0.27
	30	30	0.94	0.98	0.79	0.74	1.01	0.22
0.3	5	5	0.91	0.98	0.34	0.31	1.15	1.49
	5	10	0.88	0.96	0.37	0.33	0.63	5.67
	10	10	0.88	0.96	0.47	0.39	1.15	0.59
	20	20	0.95	0.96	0.63	0.56	1.10	0.29
	25	25	0.92	0.95	0.66	0.58	1.05	0.31
	30	30	0.88	0.92	0.77	0.65	1.03	0.18
0.4	5	5	0.87	0.96	0.31	0.23	0.21	10.65
	5	10	0.91	0.98	0.38	0.34	-0.38	14.80
	10	10	0.95	0.98	0.43	0.42	1.12	0.52
	20	20	0.95	0.95	0.59	0.53	1.08	0.30
	25	25	0.93	0.96	0.65	0.58	1.00	0.25
	30	30	0.90	0.94	0.73	0.61	1.04	0.19
0.5	5	5	0.95	0.97	0.29	0.27	0.37	5.58
	5	10	0.98	0.97	0.40	0.39	1.53	2.90
	10	10	0.90	0.95	0.33	0.27	1.11	0.49
	20	20	0.99	0.95	0.63	0.58	1.06	0.24
	25	25	0.97	0.96	0.59	0.56	1.07	0.25
	30	30	0.94	0.99	0.67	0.64	1.03	0.22

Table 3

The estimates of the logarithm of the relative potency and probabilities of testing hypotheses with $r = 1.5$

q	n_S	n_T	i) H_β^0	ii) H_μ^0	iii) $ \mu - \hat{\mu} < 0.2$	$i \wedge ii \wedge iii$	$\hat{\mu}$	S_β
0.1	5	5	0.91	0.96	0.37	0.34	1.59	3.13
	5	10	0.87	0.94	0.43	0.35	1.19	0.57
	10	10	0.93	0.92	0.50	0.45	1.10	0.39
	20	20	0.95	0.95	0.71	0.65	1.04	0.23
	25	25	0.93	0.95	0.76	0.67	1.04	0.20
	30	30	0.93	0.92	0.83	0.74	1.00	0.16
0.2	5	5	0.94	0.94	0.32	0.28	7.29	59.51
	5	10	0.90	0.91	0.35	0.29	1.15	0.51
	10	10	0.96	0.96	0.39	0.31	1.23	1.15
	20	20	0.95	0.94	0.66	0.60	1.05	0.24
	25	25	0.95	0.89	0.73	0.63	1.02	0.30
	30	30	0.93	0.95	0.76	0.68	1.01	0.23
0.3	5	5	0.93	0.95	0.28	0.26	2.87	15.87
	5	10	0.88	0.95	0.33	0.29	1.35	3.18
	10	10	0.90	0.93	0.46	0.39	1.18	0.64
	20	20	0.94	0.95	0.60	0.53	1.10	0.33
	25	25	0.93	0.95	0.65	0.59	1.06	0.35
	30	30	0.88	0.92	0.77	0.65	1.03	0.18
0.4	5	5	0.89	0.96	0.30	0.24	0.69	5.27
	5	10	0.91	0.97	0.35	0.31	1.40	3.09
	10	10	0.95	0.98	0.45	0.42	1.15	0.65
	20	20	0.94	0.95	0.60	0.54	1.09	0.34
	25	25	0.93	0.94	0.64	0.56	1.02	0.27
	30	30	0.92	0.91	0.74	0.64	1.03	0.19
0.5	5	5	0.96	0.93	0.26	0.22	0.85	2.06
	5	10	0.96	0.97	0.39	0.37	1.31	1.17
	10	10	0.90	0.98	0.36	0.30	1.12	0.54
	20	20	0.99	0.93	0.61	0.56	1.06	0.26
	25	25	0.96	0.96	0.58	0.54	1.08	0.28
	30	30	0.93	0.97	0.67	0.64	1.03	0.24

Table 4

The estimates of the logarithm of the relative potency and probabilities of acceptance of the hypotheses with $r = 2.0$

q	n_S	n_T	i) H_β^0	ii) H_μ^0	iii) $ \mu - \hat{\mu} < 0.2$	$i \wedge ii \wedge iii$	$\hat{\mu}$	$S_{\hat{\mu}}$
0.1	5	5	0.92	0.96	0.44	0.38	1.40	1.57
	5	10	0.89	0.94	0.43	0.37	1.19	0.59
	10	10	0.93	0.92	0.50	0.45	1.10	0.39
	20	20	0.95	0.96	0.70	0.64	1.04	0.24
	25	25	0.93	0.94	0.77	0.67	1.04	0.21
	30	30	0.92	0.92	0.82	0.73	1.00	0.16
0.2	5	5	0.92	0.95	0.27	0.23	0.80	6.73
	5	10	0.90	0.93	0.33	0.28	1.16	0.56
	10	10	0.91	0.95	0.42	0.33	1.26	1.37
	20	20	0.95	0.95	0.66	0.60	1.05	0.25
	25	25	0.95	0.90	0.67	0.58	1.02	0.32
	30	30	0.92	0.95	0.74	0.65	1.01	0.23
0.3	5	5	0.86	0.97	0.34	0.29	1.12	1.39
	5	10	0.88	0.95	0.32	0.27	1.86	6.15
	10	10	0.90	0.92	0.46	0.38	1.19	0.68
	20	20	0.95	0.95	0.58	0.52	1.11	0.35
	25	25	0.94	0.96	0.63	0.57	1.07	0.37
	30	30	0.91	0.90	0.73	0.60	1.03	0.19
0.4	5	5	0.91	0.92	0.30	0.25	-5.53	65.95
	5	10	0.91	0.97	0.32	0.28	1.30	2.07
	10	10	0.94	0.97	0.44	0.40	1.17	0.74
	20	20	0.95	0.95	0.57	0.53	1.10	0.36
	25	25	0.93	0.94	0.61	0.53	1.02	0.29
	30	30	0.91	0.91	0.74	0.64	1.04	0.21
0.5	5	5	0.92	0.97	0.28	0.25	-0.24	16.18
	5	10	0.95	0.97	0.37	0.35	1.32	1.19
	10	10	0.90	0.97	0.35	0.29	1.13	0.57
	20	20	0.99	0.94	0.61	0.56	1.07	0.27
	25	25	0.96	0.97	0.57	0.53	1.08	0.30
	30	30	0.93	0.97	0.63	0.601.03	0.24	

Table 5

The estimates of the logarithm of the relative potency and probabilities of acceptance of the hypotheses with $r = 10.0$

q	n_S	n_T	i) H_β^0	ii) H_μ^0	iii) $ \mu - \hat{\mu} < 0.2$	$i \wedge ii \wedge iii$	$\hat{\mu}$	$S_{\hat{\mu}}$
0.1	5	5	0.95	0.95	0.36	0.34	0.63	6.79
	5	10	0.90	0.94	0.43	0.39	1.28	0.83
	10	10	0.95	0.91	0.48	0.43	1.15	0.54
	20	20	0.94	0.95	0.61	0.56	1.08	0.37
	25	25	0.93	0.96	0.67	0.59	1.04	0.23
	30	30	0.92	0.93	0.75	0.66	1.00	0.18
0.2	5	5	0.93	0.95	0.33	0.28	2.71	8.99
	5	10	0.93	0.94	0.31	0.27	1.30	1.21
	10	10	0.93	0.96	0.37	0.28	0.87	3.68
	20	20	0.98	0.95	0.56	0.51	1.09	0.48
	25	25	0.99	0.88	0.55	0.48	1.04	0.50
	30	30	0.92	0.94	0.64	0.56	1.02	0.30
0.3	5	5	0.95	0.94	0.35	0.31	2.06	5.14
	5	10	0.94	0.95	0.30	0.26	0.69	4.86
	10	10	0.91	0.91	0.32	0.28	1.04	1.77
	20	20	0.96	0.98	0.47	0.43	1.50	3.77
	25	25	0.93	0.99	0.50	0.44	1.16	0.70
	30	30	0.93	0.92	0.60	0.52	1.04	0.27
0.4	5	5	0.86	0.97	0.25	0.19	1.47	2.70
	5	10	0.92	0.96	0.28	0.26	1.26	1.99
	10	10	0.94	0.97	0.38	0.34	1.63	6.97
	20	20	0.99	0.95	0.41	0.39	1.24	0.83
	25	25	0.92	0.92	0.47	0.38	1.10	0.49
	30	30	0.87	0.91	0.60	0.48	1.06	0.32
0.5	5	5	0.93	0.95	0.22	0.22	1.28	3.98
	5	10	0.90	0.96	0.26	0.22	4.82	18.59
	10	10	0.92	0.96	0.23	0.20	1.05	2.59
	20	20	0.98	0.95	0.49	0.36	1.14	0.50
	25	25	0.96	0.95	0.36	0.33	1.17	0.63
	30	30	0.92	0.93	0.44	0.42	1.05	0.35

The outcomes for ten percent of contamination, enclosed in Tab. 2, 3, 4 and 5 concerning the joint probability of the truthfulness of the hypotheses H_β^0 and H_μ^0 and that the condition $|\mu - \hat{\mu}| < 0.2$ is satisfied (seventh column) are illustrated on Fig. 1. From this figure we conclude that the joint probability depends heavily on the dose replications but is almost the same for the different multiplier r of the covariance matrix. By analogy, on Fig. 2 the estimates of the logarithm of the relative potency is plotted.

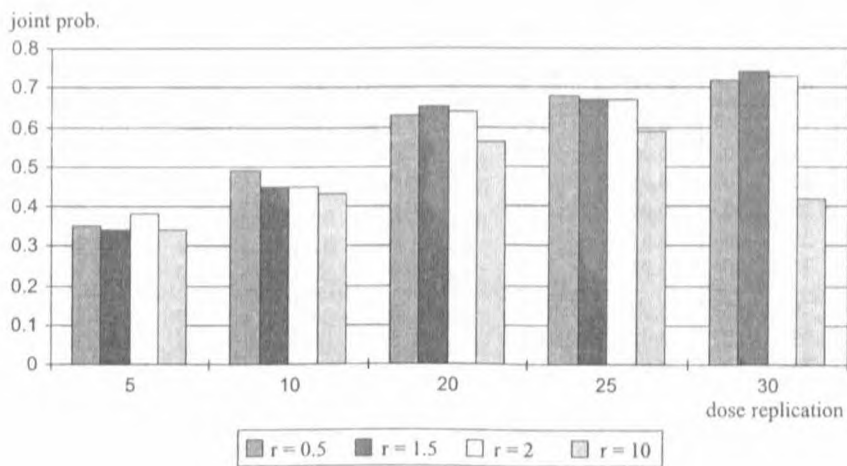


Fig. 1. Joint probability of the truthfulness of H_μ^0 and H_μ^0 and $|\mu - \hat{\mu}| < 0.2$

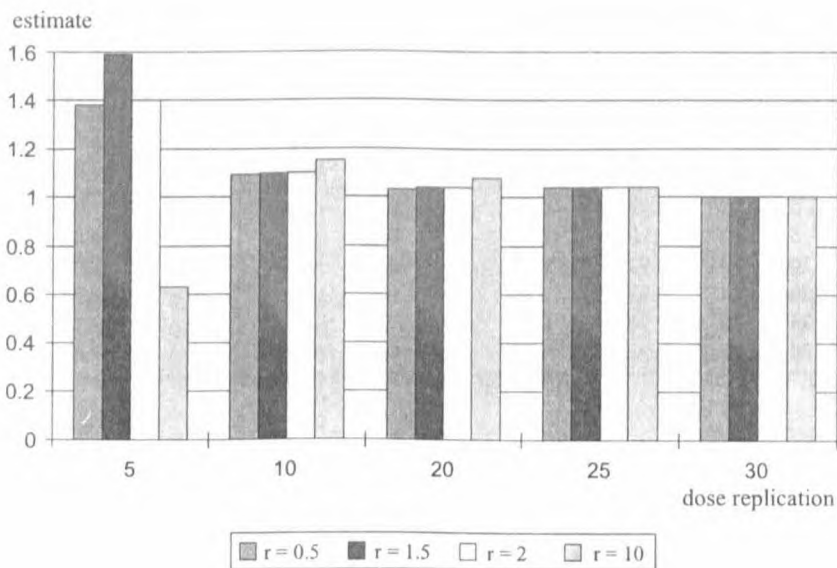


Fig. 2. Estimates of the logarithm of the relative potency for 10 percent of contaminated data

Regarding the percent q of contaminated responses equals 10, 20, 30, 40 and 50 per cent, the joint probability is plotted in Fig. 3. This figure shows, that the joint probabilities decrease a little bit where the percents of contaminated responses increase. This probability depends most on the dose replications for each q similarly as in Fig. 1. The estimates of the logarithm of the relative potency obtained for the different r and q enclosed in the penultimate column in Tab. 2, 3, 4 and 5 are illustrated in Fig. 4.

Fig. 4 shows that the estimates are very far from the true value only for the lowest number of dose replications but for the dose replications greater or equal to 10, the estimates of the logarithm of the relative potency are very close to the real value of the parameter. The estimates do not depend too much on the value of r , the multiplier of the covariance matrix, and the proportion q of the contaminated data.

joint prob.

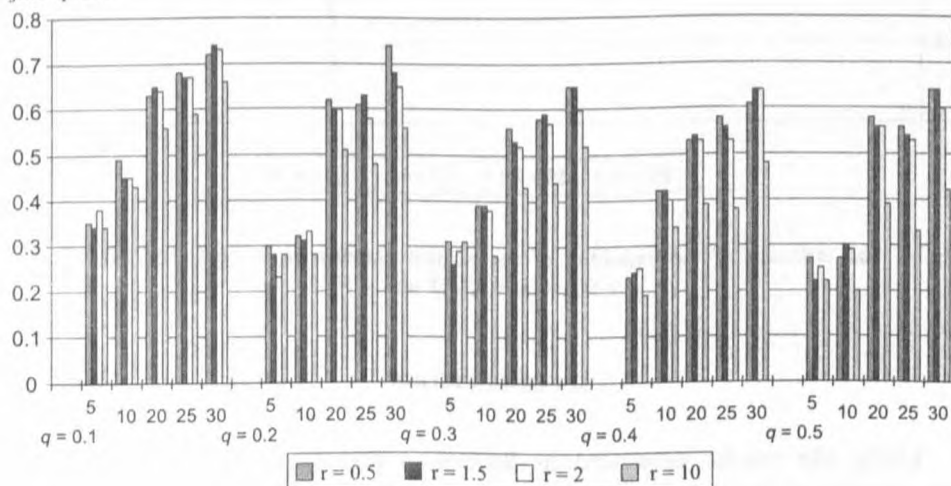


Fig. 3. Joint probability of the truthfulness of H_β^0 and H_μ^0 and $|\mu - \hat{\mu}| < 0.2$ for $q = 0.1, 0.2, 0.3, 0.4$ and 0.5

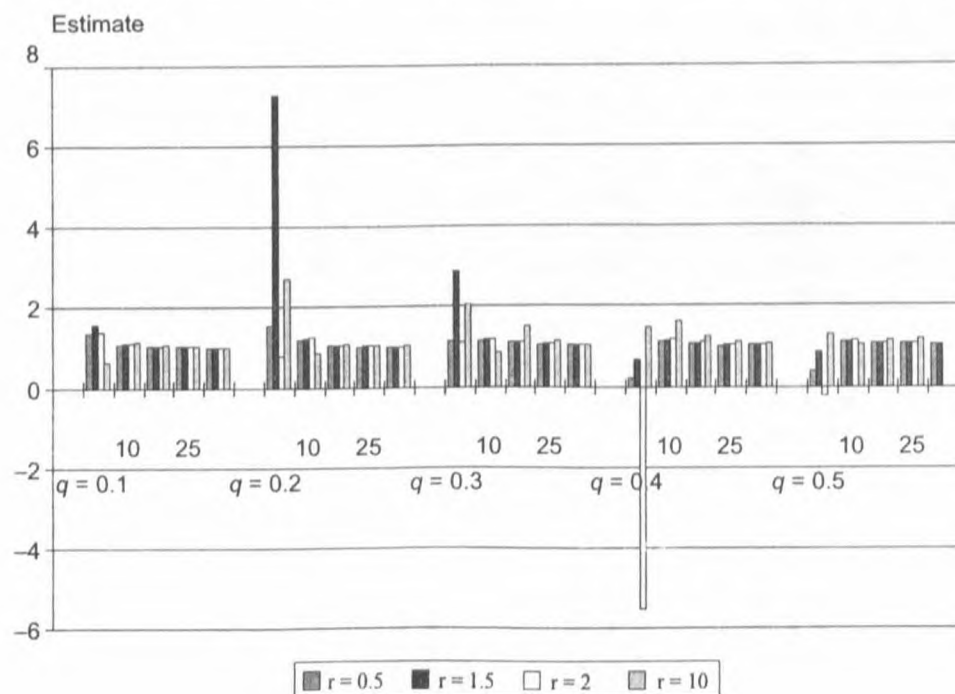


Fig. 4. The estimates of the logarithm of the relative potency for $r = 0.5, 1.5, 2.0, 10$ and $q = 0.1, 0.2, 0.3, 0.4$ and 0.5

5. CONCLUSIONS

Using the results presented in Section 3 and 4 we can conclude that the number of dose replications has the greatest influence on the logarithm of the relative potency. In particular, in the cases where responses do not have to be normally distributed, the experimenters should remember about it. With only a few dose replications, the estimates are far from the true value of the parameter, and standard deviations of the estimates are big, so with the same problem one could obtain the estimates which would differ a lot one from another. As far as the contaminated normal distribution is regarded, Tab. 2, 3, 4, 5 as well as Fig. 1, 2, 3 and 4 show that the multiplier of the covariance matrix has a very small influence on the estimator of the logarithm of the relative potency. The percent of the contaminated data sets influences the estimates but not in an essential way. Summarizing, the experiments where doses of the preparations are applied to many units, give a good estimate of the relative potency, even when the responses are not exactly of the normal distribution.

REFERENCES

- Finney D. J. (1978), *Statistical Methods in Biological Assays*, Macmillan Publishing Co., New York.
- Hanusz Z. (1995), *Relative Potency of Two Preparations in Two-way Elimination of heterogeneity Designs with Multivariate Responses*, "Biometrics", **51**, 1133-1139.
- Hanusz Z. (1998), *Influence of the Model Parameters for the Relative Potency Estimator in Multivariate Setting*, XXVIII Colloquium Biometryczne, 158-168.
- Williams D. A. (1988), *An Exact Confidence Interval for the Relative Potency Estimated from a Multivariate Bioassay*, "Biometrics", **44**, 861-867.