⟨Research Paper⟩

Journal of the Korean Statistical Society
Vol. 16, No. 2, 1987

Chances of Simpson's Paradox+

Jong Woo Jeon*, Han Yeong Chung* and Jong Sig Bae*

ABSTRACT

The Simpson's paradox is a paradoxical phenomenon which might occur when analyzing 2×2 contingency table. This paper considers the role of probability assignment of the experimental units in reducing the chances of Simpson's paradox. Numerical results are given to illustrate how the chance of Simpson's paradox behaves.

1. Introduction

When a 2×2 contingency table is analyzed, one might face a paradoxical phenomenon. For example, consider the data in Table 1.1 where 40 patients were given a treatment, T, and 40 assigned to a control, $\bar{T}$. The patients either recovered, R, or did not, $\bar{R}$.

Table 1.1 Recovery rate under treatment and control

	R	$\bar{R}$		Recovery rate
$\bar{T}$	20	20	40	50%
T	16	24	40	40%
	36	44	80	

The table shows that the recovery rate for control is 10% higher than that for the treatment. Now, if the sex of the patients is considered as an additional variable, the data can be classified into the two sets of tables, one for the male patients, M, and the other for the females, $\bar{M}$.

+ Research supported in part by MOE through the Research Institute for Basic Sciences, SNU, 1986~1987.

* Department of Computer Science and Statistics, SNU

Table 1.2 Recovery rate under treatment and control with sex as an added variable

Males(M)	R	$\bar{R}$		Recovery rate
$\bar{T}$	18	12	30	60%
T	7	3	10	70%
	25	15	40	

Females($\bar{M}$)	R	$\bar{R}$		Recovery rate
$\bar{T}$	2	8	10	20%
T	9	21	30	30%
	11	29	40	

It is then possible, as is shown in Table 1.2, that the recovery rate for the control is lower than that for the treatment, both for the males and for the females.

In the literature, this phenomenon is referred to as Simpson's pardox, Simpson(1951). The recent references in Simpson's paradox are Blyth (1972), Lindley and Novick (1981), Wagner (1982), Shapiro (1982), and Paik (1985) among others.

To explain this phenomenon more formally, let $R, \bar{R}, T, \bar{T}, M$, and $\bar{M}$ be the events with obvious meanings. Then,

$$P(R|T)=P(M|T)P(R|TM)+P(\bar{M}|T)P(R|T\bar{M})$$

and

$$P(R|\bar{T})=P(M|\bar{T})P(R|\bar{T}M)+P(\bar{M}|\bar{T})P(R|\bar{T}\bar{M}).$$

A pardoxical phenomenon is the situation where

$$P(R|T)<P(R|\bar{T}) \text{ but}$$
$$P(R|TM)>P(R|\bar{T}M) \text{ and } P(R|T\bar{M})>P(R|\bar{T}\bar{M}).$$

Assume that

$$\alpha=P(R|TM)-P(R|\bar{T}M)=P(R|T\bar{M})-P(R|\bar{T}\bar{M})$$
$$\delta=P(R|TM)-P(R|T\bar{M})=P(R|\bar{T}M)-P(R|\bar{T}\bar{M}),$$

where α and $\delta (0<\alpha<\delta<1)$ are positive constants which might be interpreted as the treatment effect and the sex effect on the recovery rate respectively.

Recently, under the above set up, Huh(1987) examined the chance of Simpson's pardox when each of patients is assigned randomly to the treatment or the control. That is, each patient (male or famale) is assigned independently to the treatment T with probability $\frac{1}{2}$ or to the control $\bar{T}$ with probability $\frac{1}{2}$. He concluded that the chance decreases as the sample size increases and that, for given sample size, the random assignment works efficiently to prevent Simpson's paradox only when δ/α, the ratio of the sex effect over the treatment effect, is not large.

In the present study, we consider the chance of Simpson's paradox under different probabilities of assignment. We shall deal with two cases. One is the unblanced case where the male and the female patients are assigned independently to the treatment with probability p and $1-p$ respectively. The other is the balanced case in which all the patients are assigned independently to the treatment with equal probability p.

For the unbalanced case, a conclusion is that the chance of Simpson's paradox decreases as p increases. In other words, in order to prevent Simpson's paradox, we should assign higher probability to the subgroup with higher recovery rate. For the balanced case, conclusions are parallel to those by Huh(1987).

In Section 2, a graphical illustrations for Simpson's paradox is presented. In Section 3, we treat the chance of Simpson's paradox for the case of unbalanced assignment and study the case of balanced assignment in Section 4.

2. A Graph of Simpson's Paradox

Consider Figure 2.1. The lined circles represent the subgroup of the treatment while dotted circles do that of the control. The Y-coordinate of the center of each circle stands for the corresponding recovery rate. The area of each circle is proportional to the sample size of the subgroup. For the Figure 2.1, we refer Lindely and Novick (1981) and Paik(1985).

Since what really matters in this set up is the differences $P(R|TM)-P(R|\bar{T}M)$, $P(R|T\bar{M})-P(R|\bar{T}\bar{M})$, we can let without loss of generality $P(R|TM)=1$ and $P(R|\bar{T}\bar{M})=0$.

It is then clear from Figure 2.2 that the Simpson's paradox arises if and only if

$$\alpha+P(M|T)\delta<\delta P(M|\bar{T})$$

or

$$P(M|T)-P(M|\bar{T})<-\frac{\alpha}{\delta}. \tag{2.1}$$

The same conclusion was obtained in Huh(1987).

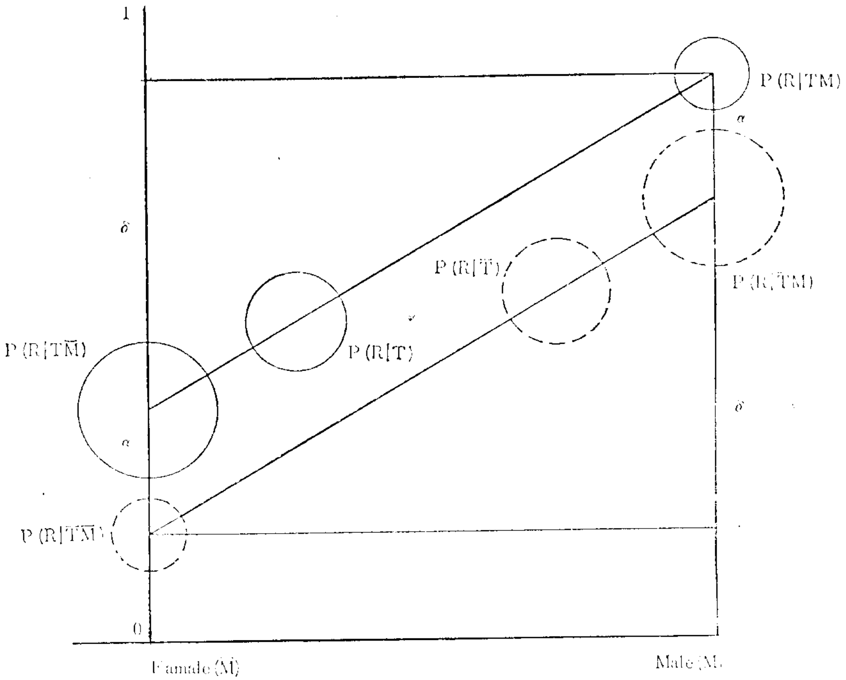

Figure 2.1 Graphical repesentation of Simpson's pardox

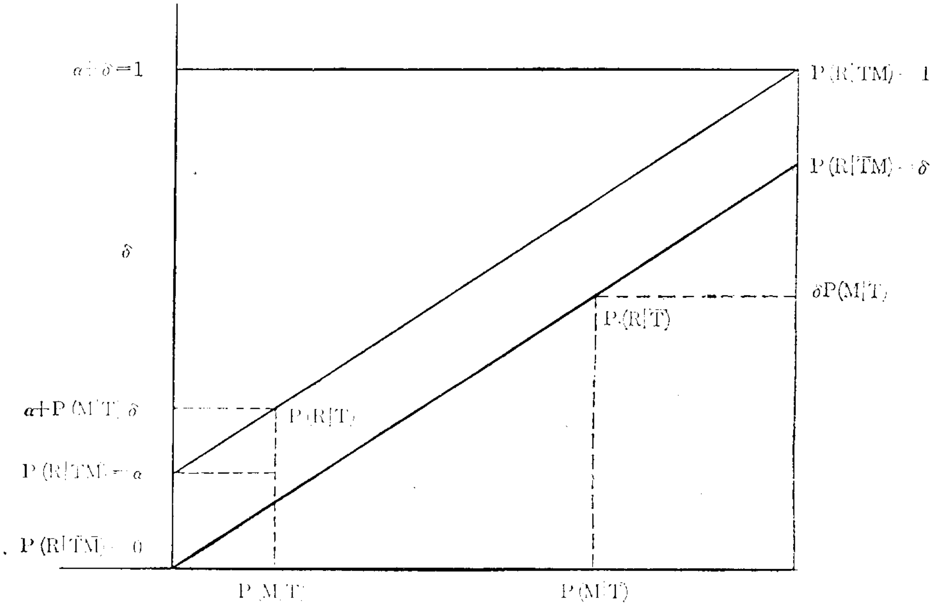

Figure 2.2 Graphical repesentation of Simpson's paradox

3. Unbalanced Case

Suppose that there are $n(=n_1+n_2)$ patients where n_1 patients are male (M) and n_2 patients are female $(\bar{M})$. Suppose also that each male patient is assigned independently to the treatment T with probability p and each female patient is assigned to the treatment T with probability $1-p$.

Let B_1 be the number of male patients assigned to the treatment and B_2 be the number of female patients assigned to the treatment. Then, B_1 and B_2 are independent binomial variables with parameters (n_1, p) and $(n_2, 1-p)$ respectively.

We are interested in the probability of

$$P(M|T)-P(M|\bar{T})<-\frac{\alpha}{\delta} \tag{3.1}$$

Let $P(n_1, n_2 : p)$ denote such probability. Then, since $P(M|T)$ and $P(M|\bar{T})$ are given by $B_1/(B_1+B_2)$ and $(n_1-B_1)/(n-B_1-B_2)$ respectively,

$$P(n_1, n_2 : p)$$
$$=P\left\{B_1/(B_1+B_2)-(n_1-B_1)/(n-B_1-B_2)<-\frac{\alpha}{\delta}\right\}. \tag{3.2}$$

Note that $P(n_2, n_1 : p)=P(n_1, n_2 : p)$. This follows from $\frac{B_1}{B_1+B_2}-\frac{n_1-B_1}{n-B_1-B_2}=\frac{n_2-B_2}{n-B_1-B_2}-\frac{B_2}{B_1+B_2}$ and the fact fact that n_2-B_2 and n_1-B_1 are independent binomial random variables with parameters (n_2, p) and $(n_1, 1-p)$ respectively.

We tabulate (3.2) in Table 3.1, for $\delta/\alpha=2, 4, 6, 8$, $n=20, 60, 100$, $n_1/n_2=1, 2, 3$, and $p=1/4$, $1/3$, $1/2$, $2/3$, $3/4$.

By inspecting Table 3.1, we see that the chance of Simpson's paradox decreases as p increases for each n, δ/α and the sex ratio n_1/n_2. Figure 3.1 shows a behavior of chance of Simpson's paradox for fixed $n=20$, $\delta/\alpha=8$ for various values of p.

Remark. Since $\frac{B_1}{n}=\frac{B_1}{n_1}\frac{n_1}{n}\to\frac{p}{2}$ w. p. 1 and $\frac{B_2}{n}\to\frac{1-p}{2}$ w. p. 1 as n approaches to the infinity,

$$\frac{B_1/n}{(B_1+B_2)/n}-\frac{(n_1-B_1)/n}{(n-B_1-B_2)/n}\to 2p-1 \text{ w. p. 1.}$$

Thus, for $p_0=\frac{1}{2}\left(1-\frac{\alpha}{\delta}\right)$, we have

$$\lim_{\substack{n\to\infty \\ n_1/n_2=1}} P(n_1, n_2 : p)=\begin{cases}1 & \text{if } p<p_0 \\ 0 & \text{if } p\geq p_0\end{cases}.$$

This fact is reflected in Figure 3.2 which shows that the slopes of the graphs become steeper as n increases.

Table 3.1 The chances of Simpson's paradox

n					$\delta/\alpha=2$	$\delta/\alpha=4$	$\delta/\alpha=6$	$\delta/\alpha=8$
20	$n_1=10$,	$n_2=10$						
			$P=$	.75	.0000	.0002	.0010	.0011
			$P=$	.67	.0002	.0039	.0131	.0147
			$P=$	.50	.0214	.1340	.2526	.2629
			$P=$	.33	.2999	.6644	.8100	.8163
			$P=$	.25	.6194	.8994	.9592	.9610
20	$n_1=13$,	$n_2=7$						
			$P=$	.75	.0000	.0004	.0012	.0019
			$P=$	.67	.0001	.0049	.0137	.0190
			$P=$	.50	.0088	.1188	.2359	.2704
			$P=$	.33	.1719	.5965	.7794	.8008
			$P=$	.25	.4394	.8529	.9455	.9515
20	$n_1=15$,	$n_2=5$						
			$P=$	.75	.0000	.0012	.0019	.0035
			$P=$	.67	.0000	.0080	.0140	.0250
			$P=$	.50	.0026	.1291	.1928	.2644
			$P=$	.33	.0862	.5787	.6896	.7557
			$P=$	.25	.2788	.8289	.8974	.9230
60	$n_1=30$,	$n_2=30$						
			$P=$	.75	.0000	.0000	.0000	.0000
			$P=$	.67	.0000	.0000	.0001	.0002
			$P=$	.50	.0001	.0266	.1226	.1832
			$P=$	.33	.1098	.7581	.9320	.9603
			$P=$	.25	.5735	.9848	.9985	.9994
60	$n_1=40$,	$n_2=20$						
			$P=$	.75	.0000	.0000	.0000	.0000
			$P=$	.67	.0000	.0000	.0001	.0002
			$P=$	.50	.0000	.0199	.0874	.1587
			$P=$	.33	.0437	.6781	.8724	.9336
			$P=$	.25	.3685	.9691	.9944	.9982
60	$n_1=45$,	$n_2=15$						
			$P=$	.75	.0000	.0000	.0000	.0000
			$P=$	.67	.0000	.0000	.0001	.0003
			$P=$	.50	.0000	.0155	.0708	.1342
			$P=$	.33	.0131	.5722	.8062	.8880
			$P=$	.25	.1897	.9346	.9855	.9942

100	$n_1=50$,	$n_2=50$						
			$P=$	.75	.0000	.0000	.0000	.0000
			$P=$	.67	.0000	.0000	.0000	.0000
			$P=$	.50	.0000	.0061	.0449	.0977
			$P=$	.33	.0462	.8135	.9570	.9833
			$P=$	.25	.5546	.9973	.9999	1.0000
100	$n_1=67$,	$n_2=33$						
			$P=$	.75	.0000	.0000	.0000	.0000
			$P=$	.67	.0000	.0000	.0000	.0000
			$P=$	.50	.0000	.0037	.0418	.1012
			$P=$	.33	.0128	.7100	.9314	.9742
			$P=$	.25	.3227	.9909	.9995	.9999
100	$n_1=75$,	$n_2=25$						
			$P=$	.75	.0000	.0000	.0000	.0000
			$P=$	.67	.0000	.0000	.0000	.0000
			$P=$	.50	.0000	.0021	.0283	.0749
			$P=$	.33	.0017	.5454	.8577	.9376
			$P=$	.25	.1180	.9638	.9971	.9993

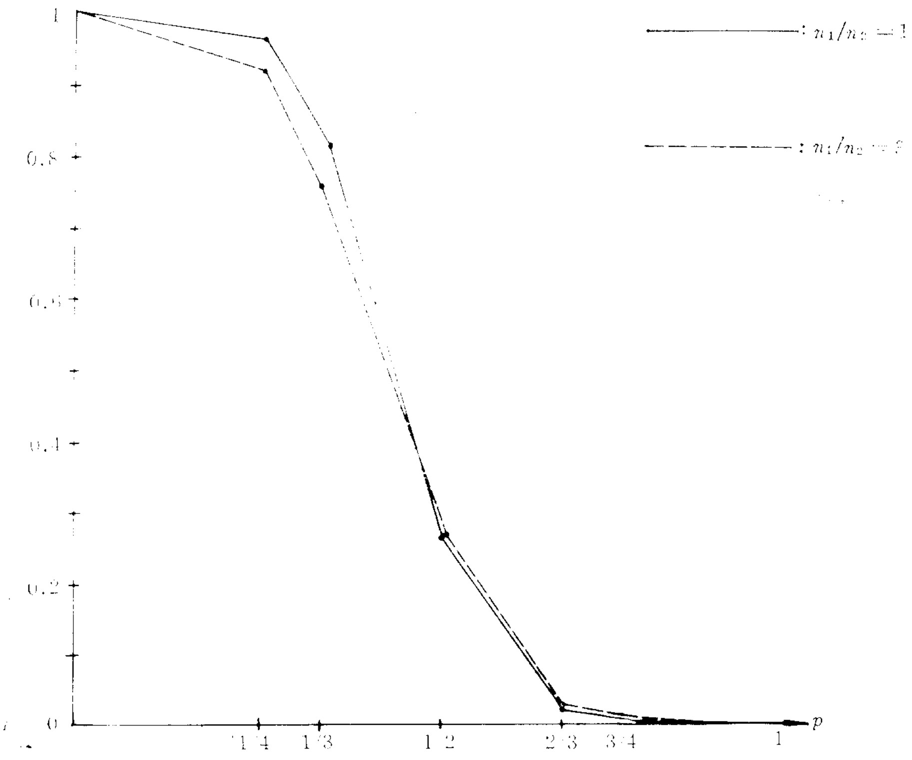

Figure 3.1 Behavior of chance of Simpson's paradox: $n=20$ and $\delta/\alpha=8$

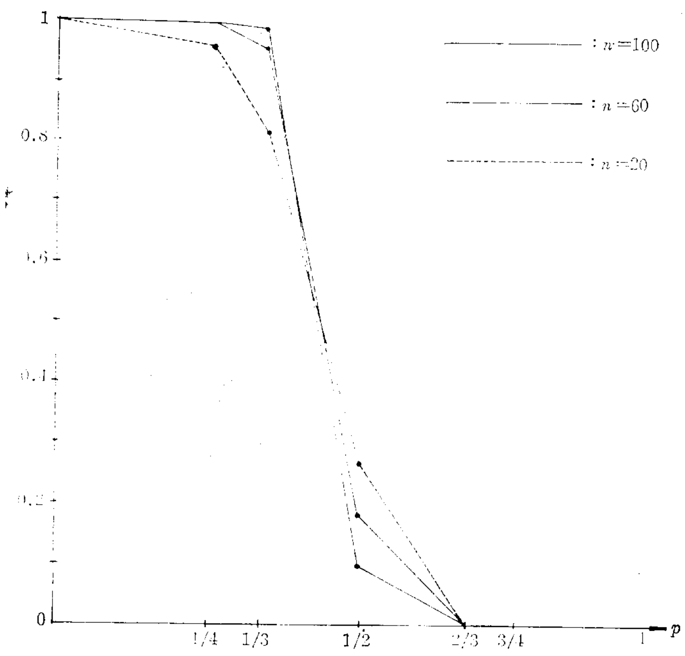

Figure 3.2 Behavior of chance of Simpson's paradox: $n_1/n_2=1$ and $\delta/\alpha=8$

4. Balanced Case

In this section, we assume that each patient, either male or female, is assigned independently to the teatment T with probability p or to the control $\bar{T}$ with probability $1-p$.

Let B_1 and B_2 be again the number of males and females assigned to the treatment respectively.

That is, B_i, $i=1,2$, are independent and distributed as Binomial (n_i, p), $i=1,2$, respectively.

Let $Q(n_1, n_2 : p)$ be the probability that (3.1) occurs. Then, as before,

$$Q(n_1, n_2 : p)=P\left\{B_1/(B_1+B_2)-(n_1-B_1)/(n-B_1-B_1)<-\frac{\alpha}{\delta}\right\}. \qquad (4.1)$$

Note in this case that $Q(n_2, n_1 : 1-p)=Q(n_1, n_2 : p)$ by the similar reasons as in Section 3.

The values of $Q(\mathrm{n}_1, \mathrm{n}_2 : p)$ are tabulated in Table 4.1.

By inspecting Table 4.1, we see that the chance of Simpson's paradox decreases to zero as n increases for each $\delta/\alpha, p$, and sex ratio n_1/n_2. We also note that the value of p has little effects on the chance of Simpson's paradox. The figure 4.1 shows the behavior of the chance for various n and p when $\delta/\alpha=6$, $n_1/n_2=1$.

Table 4.1 The chance of Simpson's paradox

n					$\delta/\alpha=2$	$\delta/\alpha=4$	$\delta/\alpha=6$	$\delta/\alpha=8$
20	$n_1=10$,	$n_2=10$						
			$P=$	.75	.0579	.1697	.2808	.3513
			$P=$	.67	.0285	.1443	.2565	.3073
			$P=$	.50	.0214	.1340	.2526	.2629
			$P=$	.33	.0285	.1443	.2565	.3073
			$P=$	.25	.0579	.1697	.2808	.3513
20	$n_1=31$,	$n_2=7$						
			$P=$	.75	.0033	.1963	.2684	.2858
			$P=$	.67	.0065	.1611	.2379	.2803
			$P=$	.50	.0088	.1188	.2359	.2704
			$P=$	.33	.0167	.1256	.2647	.2741
			$P=$	.25	.0309	.1501	.2777	.3100
20	$n_1=15$,	$n_2=5$						
			$P=$	.75	.0000	.2410	.2567	.2939
			$P=$	.67	.0002	.1610	.2011	.2672
			$P=$	.50	.0026	.1291	.1928	.2644
			$P=$	.33	.0096	.1407	.2211	.2514
			$P=$	.25	.0200	.1697	.2438	.2519
60	$n_1=30$,	$n_2=30$						
			$P=$	.75	.0006	.0495	.1471	.2016
			$P=$	.67	.0002	.0368	.1215	.1794
			$P=$	.50	.0001	.0266	.1226	.1832
			$P=$	.33	.0002	.0368	.1215	.1794
			$P=$	.25	.0006	.0495	.1471	.2016
60	$n_1=40$,	$n_2=20$						
			$P=$	.75	.0000	.0363	.1218	.1999
			$P=$	.67	.0000	.0264	.1049	.1645
			$P=$	.50	.0000	.0199	.0874	.1587
			$P=$	.33	.0001	.0294	.1012	.1718
			$P=$	.25	.0005	.0489	.1205	.1875
60	$n_1=45$,	$n_2=15$						
			$P=$	.75	.0000	.0228	.1043	.1658
			$P=$	.67	.0000	.0172	.0817	.1451
			$P=$	.50	.0000	.0155	.0708	.1342
			$P=$	.33	.0001	.0261	.0931	.1524
			$P=$	.25	.0003	.0393	.1195	.1727

100	n_1=50,	n_2=50					
			P= .75	.0000	.0158	.0757	.1424
			P= .67	.0000	.0093	.0598	.1231
			P= .50	.0000	.0061	.0449	.0977
			P= .33	.0000	.0093	.0598	.1231
			P= .25	.0000	.0158	.0757	.1424
100	n_1=67,	n_2=33					
			P= .75	.0000	.0090	.0624	.1254
			P= .67	.0000	.0054	.0497	.1021
			P= .50	.0000	.0037	.0418	.1012
			P= .33	.0000	.0069	.0509	.1080
			P= .25	.0000	.0125	.0686	.1305
100	n_1=75,	n_2=25					
			P= .75	.0000	.0039	.0447	.1048
			P= .67	.0000	.0023	.0311	.0875
			P= .50	.0000	.0021	.0283	.0749
			P= .33	.0000	.0040	.0370	.0881
			P= .25	.0000	.0088	.0493	.1122

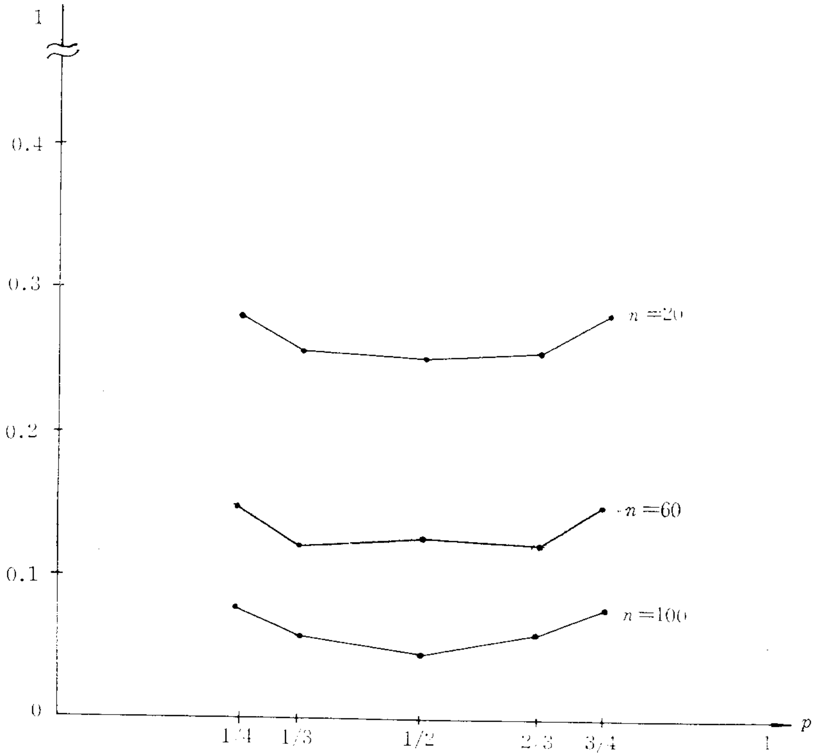

Figure 4.1 Behavior of the chance of Simpson's paradox: $\delta/\alpha=\overline{6}$, $\overline{n_1/n_2}=1$

References

(1) Blyth, C.R.(1972). On Simpson's Paradox and Surething Principle, *Journal of the American Statistical Association*. 67, 364~366.

(2) Huh, M.H.(1987). Simpson's Paradox and Randomization, *Journal of the Korean Statistical Society*, 16, 113~116.

(3) Lindley, D.V. and Novick, M.R.(1981). The Role of Exchangeability in Inference, *The Annals of Statistics*, 9, 45~58.

(4) Paik, M.J.(1985). A Graphic Representation of a Three-way Contingency Tables: Simpson's Pardox and Correlation, *The American Statistician*, 39, 53~54.

(5) Shapiro, S.H.(1982). Collapsing Contingency Tables, A Geometric Approach, *The American Statistician*, 36, 43~46.

(6) Simpson, E.H.(1951). The Interpretation of Interaction in Contingency Tables, *Journal of the Royal Statistical Society Series B*. 13, 238~241.

(7) Wagner, C.H.(1982). Simpson's Paradox in Real Life, *The American Statistician*, 36, 46~48.