

Improved Sib-Pair Linkage Test for Disease Susceptibility Loci

Julian J. Faraway

Department of Statistics, University of Michigan, Ann Arbor

An improved sib-pair test for linkage is introduced which is superior to the previously proposed tests. The test is derived from the standard chi-squared goodness of fit statistic by restricting the alternative hypothesis to the genetically possible. Critical values are given and exact power comparisons are made with the previously proposed tests. The new test is shown to be more powerful for finite samples as well as being asymptotically uniformly most powerful. © 1993 Wiley-Liss, Inc.

Key words: most powerful test, sib-pair method, chi-squared goodness of fit statistic

INTRODUCTION

Sib-pair methods for detecting linkage are useful because they make no assumptions about the specific mode of inheritance of the disease under study. Linkage studies using this method have become more popular recently and promise to become more so with the increasing availability of sufficiently polymorphic markers. In the past, Cudworth and Woodrow [1975], de Vries et al. [1976], and Weitkamp et al. [1981] all used sib-pair methods. More recently, several studies have used these methods, among these being Concannon et al. [1990], Cox et al. [1988], Fimmers et al. [1989], Forman et al. [1992], and Jeunemaitre et al. [1992].

The test we propose is based on data from pairs of siblings, who are both affected with the disease under study and who are from different families. Given information on the number of marker locus alleles that the sibs share identical by descent (IBD), we can test for linkage. Several tests for linkage have been proposed previously. A test based on the proportion of sib pairs with two marker alleles IBD, termed the "two-allele test," was proposed by Day and Simons [1976] and by Suarez et al. [1978] and used in Weitkamp et al. [1981]. Another test based on the mean number of marker alleles IBD, termed the "mean test," was suggested by de Vries et al. [1976] and

Received for publication July 21, 1992; revision accepted April 30, 1993.

Address reprint requests to Julian J. Faraway, Department of Statistics, University of Michigan, Ann Arbor, MI 48109.

Green and Woodrow [1977]. Weitkamp et al. [1981] also used a chi-squared “goodness of fit” test. Blackwelder and Elston [1985] compared the powers of these tests and concluded that the goodness of fit was the worst and that the mean test was usually more powerful than the two-allele test. Schaid and Nick [1990] proposed using a linear combination of the number of marker alleles IBD as a test statistic, but optimizing this statistic requires knowledge of the parameters specifying the inheritance. They also proposed the use of the maximum of the mean and the two-allele statistics and demonstrated that its power was close to the best of these two across a range of parameters. In relation to this paper see also the letter to the editor by Knapp [1991] and reply by Schaid and Nick [1991]. The work of Fishman et al. [1978] discusses a more general test statistic, which does not depend on the observed IBD distribution. Majumber and Pal [1987] discuss related tests when parental haplotype information is available.

The test we introduce here is derived by restricting the alternative hypothesis to genetically possible values and then computing the appropriate test statistic. It should be emphasized that this restriction assumes that the disease is due to a single major locus. We have concentrated on the case where both sibs are affected, because this is the most powerful sampling scheme and the one most likely to be used in practice. It would, however, be possible to adapt the test statistic for data involving singly affected or unaffected sib pairs.

OLD TEST STATISTICS

Let p_0, p_1, p_2 be the probabilities that a given affected sib-pair shares 0, 1, or 2 marker alleles IBD. If there is no linkage, these probabilities will be $1/4, 1/2, 1/4$, respectively. Let the number of pairs of sibs be n and the observed proportions of pairs with 0, 1, 2, marker alleles IBD be $\hat{p}_0, \hat{p}_1, \hat{p}_2$. The test statistic for the two-allele test is $t_1 = (\hat{p}_2 - 1/4)/\sqrt{3/1603n}$ and tests whether the observed proportion \hat{p}_2 is significantly different from the expected $1/4$ under the hypothesis of no linkage, whereas the mean test uses $t_2 = (2\hat{p}_2 + \hat{p}_1 - 1)/\sqrt{1/2n}$ as a test statistic and tests whether the mean number of marker alleles IBD, $2\hat{p}_2 + \hat{p}_1$, is significantly different from the expected 1 under the hypothesis of no linkage. Both formulations lead to one-sided tests where the asymptotic distribution of the test statistic under the null hypothesis is standard normal. Both these statistics throw away a certain amount of information and cannot be expected to be optimal. The chi-squared goodness of fit is based on the test statistic

$$Y = 4n(\hat{p}_0 - 1/4)^2 + 2n(\hat{p}_1 - 1/2)^2 + 4n(\hat{p}_2 - 1/4)^2$$

for testing the null hypothesis of $H_0 : p_0 = p_2 = 1/4, p_1 = 1/2$ against the alternative hypothesis $H_A : p_0 + p_1 + p_2 = 1$. Y then has a null distribution of a chi-squared with 2 degrees of freedom but the test only rejects the null when $t_2 > 0$ and Y exceeds the upper 2α point. However, the alternatives are restricted; (p_0, p_1, p_2) can only take values defined by the model for inheritance. It should be no surprise that this test was the least powerful in Blackwelder and Elston’s [1985] study because the alternative hypothesis includes mostly infeasible values of (p_0, p_1, p_2) . Note that the other two tests also allow infeasible values of (p_0, p_1, p_2) which would lead one to expect that they too are suboptimal.

A test statistic proposed by Schaid and Nick [1990] is based on a *linear* combina-

tion of $(\hat{p}_0, \hat{p}_1, \hat{p}_2)$, which is asymptotically most powerful test, among linear combinations, against a specified point (p_0, p_1, p_2) alternative. However, since the object of sib-pair tests is to avoid specifying a point alternative, the alternative hypothesis must range over all physically possible values of (p_0, p_1, p_2) and so, as the authors admit, this test is not practically applicable. They also proposed $\max(t_1, t_2)$ as a test statistic.

NEW TEST STATISTIC

If we restrict the alternative hypothesis to possible values of (p_0, p_1, p_2) and compute the test statistic accordingly, then we can expect a more powerful test. The procedure is explained in Lehmann [1986:pp 480 ff]. This test is an asymptotically uniformly most powerful invariant test. So as n gets large this test is, in some sense, the best possible. We shall see that it dominates the other tests for small n also.

Since the alternative hypothesis of our test allows all feasible (p_0, p_1, p_2) (except the point of the null hypothesis, $p_0 = p_2 = 1/4, p_1 = 1/2$), it is practical (in that parameters that are not usually known, do not need to be specified) as well as asymptotically at least as powerful, at any point in (p_0, p_1, p_2) , as the aforementioned test.

Now, following Suarez et al. [1978], we assume the trait locus has two alleles T and t with frequencies p and $q = 1 - p$, respectively. Let the penetrance probabilities for the three genotypes, TT, Tt, tt , be f_1, f_2, f_3 , respectively. We assume that $f_1 \geq f_2 \geq f_3$. The prevalence of the trait in the population (K) is then $K = p^2 f_1 + 2pqf_2 + q^2 f_3$ and the additive (V_A) and dominance (V_D) variances are

$$V_A = 2pq(p(f_2 - f_1) + q(f_3 - f_2))^2 \quad V_D = p^2 q^2 (f_1 - 2f_2 + f_3)^2.$$

Let the recombination fraction be θ and $\varphi = \theta + (1 - \theta)^2$, then the values of (p_0, p_1, p_2) are then determined by K, V_A, V_D , and θ as described in table 3 of Suarez et al. [1978]:

$$p_0 = 1/4 - \frac{(\varphi - 1/2)V_A + (2\varphi - \varphi^2 - 3/4)V_D}{4(K^2 + V_A/2 + V_D/4)}$$

$$p_1 = 1/2 - \frac{2(\varphi^2 - \varphi + 1/4)V_D}{4(K^2 + V_A/2 + V_D/4)}$$

$$p_2 = 1/4 + \frac{(\varphi - 1/2)V_A + (\varphi^2 - 1/4)V_D}{4(K^2 + V_A/2 + V_D/4)}$$

(p_0, p_1, p_2) can only take values in a region F defined by $0 \leq f_1, f_2, f_3, p \leq 1, f_1 \geq f_2 \geq f_3, 0 \leq \theta \leq 1/2$. The region is determined by three linear constraints. The first, $p_1 + p_2 \leq 1$, follows trivially from the fact that $p_0 + p_1 + p_2 = 1$. The second, $p_1 \leq 1/2$, can be derived easily from the formula for p_1 given above. The third restriction, $3p_1/2 + p_2 \geq 1$, is more difficult to obtain. The triangle F is equivalent to that found independently by Holmans [1993]. In addition, it may be shown that this triangle cannot be made any smaller, which indicates that the power of the test cannot be improved further without the imposition of tighter constraints on the parameters. See the Appendix for details.

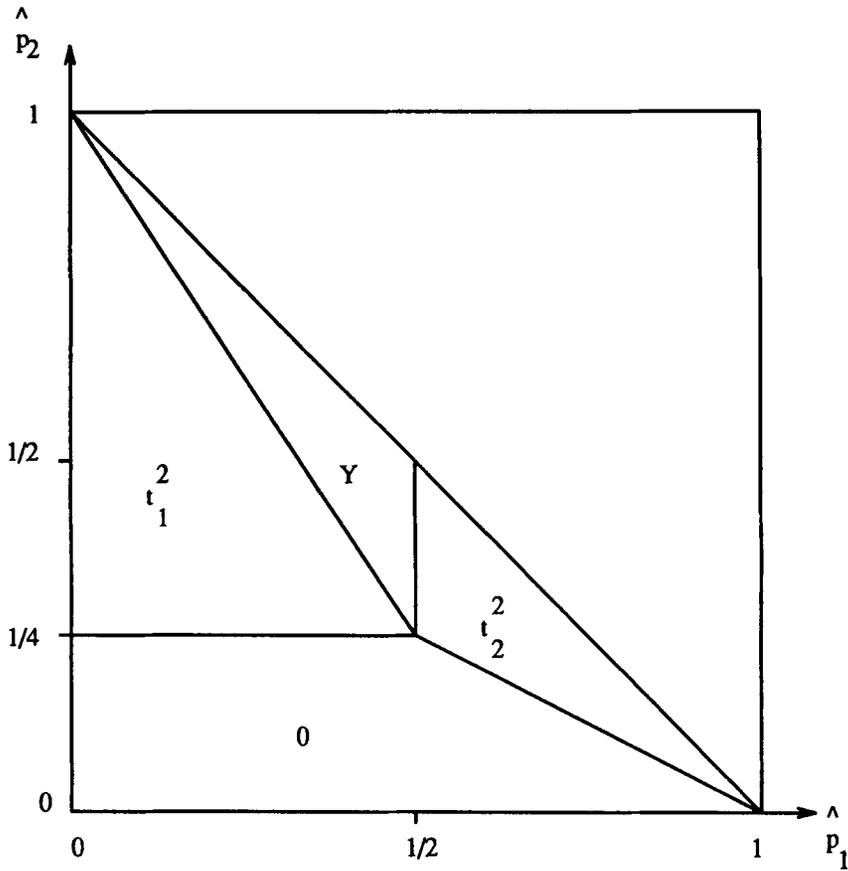


Fig. 1. The value of the new test statistic depends on the observed value of \hat{p}_1 and \hat{p}_2 , shown in this diagram.

Suarez et al. [1976] determined the feasible region for (V_A, V_D) for fixed values of K , but this is not sufficient to determine the feasible region of (p_0, p_1, p_2) for general K .

The computation of the test statistic requires that we must find the point in F that minimizes the weighted distance to $(\hat{p}_0, \hat{p}_1, \hat{p}_2)$ and the compute the chi-squared statistic based on that point, i.e., find the (v_0, v_1, v_2) that minimize

$$S = 4(\hat{p}_0 - v_0)^2 + 2(\hat{p}_1 - v_1)^2 + 4(\hat{p}_2 - v_2)^2$$

subject to $(v_0, v_1, v_2) \in F$ and then the test statistic is

$$T = 4n(v_0 - 1/4)^2 + 2n(v_1 - 1/2)^2 + 4n(v_2 - 1/4)^2.$$

This is a straightforward quadratic programming problem. For a reference on solving such problems see Fletcher [1987]. Unless $(p_0, p_1, p_2) \in F$, S will be minimized

TABLE I. Critical Values for T

Tail probability	0.05	0.01	0.005	0.001	0.0005	0.0001
Critical value	3.42	6.48	7.71	11.00	12.60	16.22

on the boundary of F , so the test statistic may be explicitly calculated. There are four possibilities, depending on the region in which $(\hat{p}_0, \hat{p}_1, \hat{p}_2)$ falls. Serendipitously, the test statistic is a hybrid of three previously used tests.

Region	Test statistic
F	Y
$2\hat{p}_2 + \hat{p}_1 \geq 1, \hat{p}_1 \geq 1/2$	t_2^2
$3p_1/2 + p_2 < 1, \hat{p}_2 \geq 1/4$	t_1^2
$2\hat{p}_2 + \hat{p}_1 < 1, \hat{p}_2 < 1/4$	0

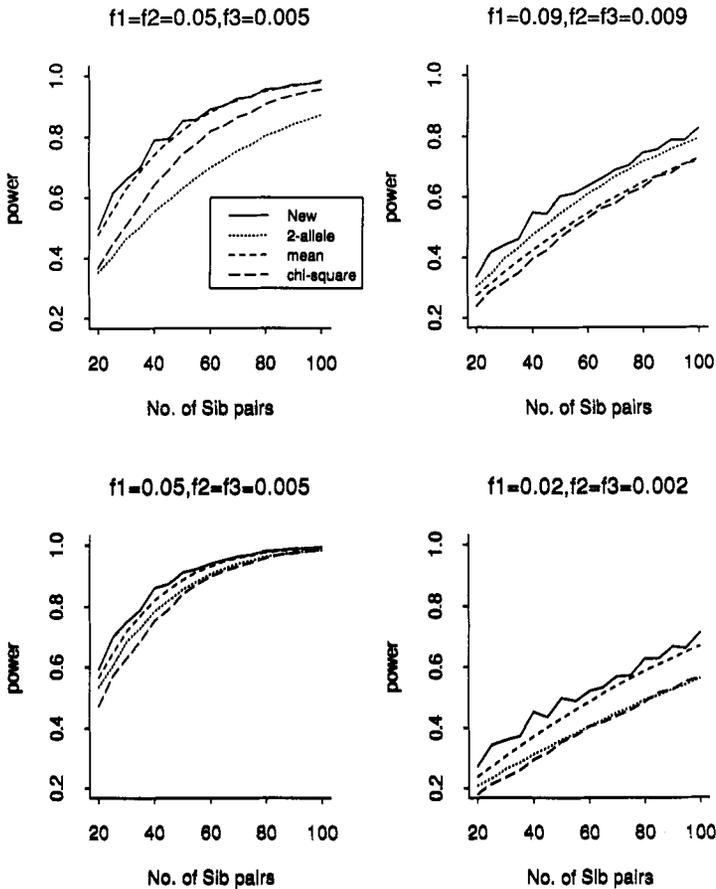


Fig. 2. Exact power for the new test (T), the two-allele test, (t_1), the mean test, (t_2), and the chi-squared goodness of fit test (Y) for a nominal 5% level of significance, complete linkage, $\theta = 0$, population frequency, $K = 0.01$.

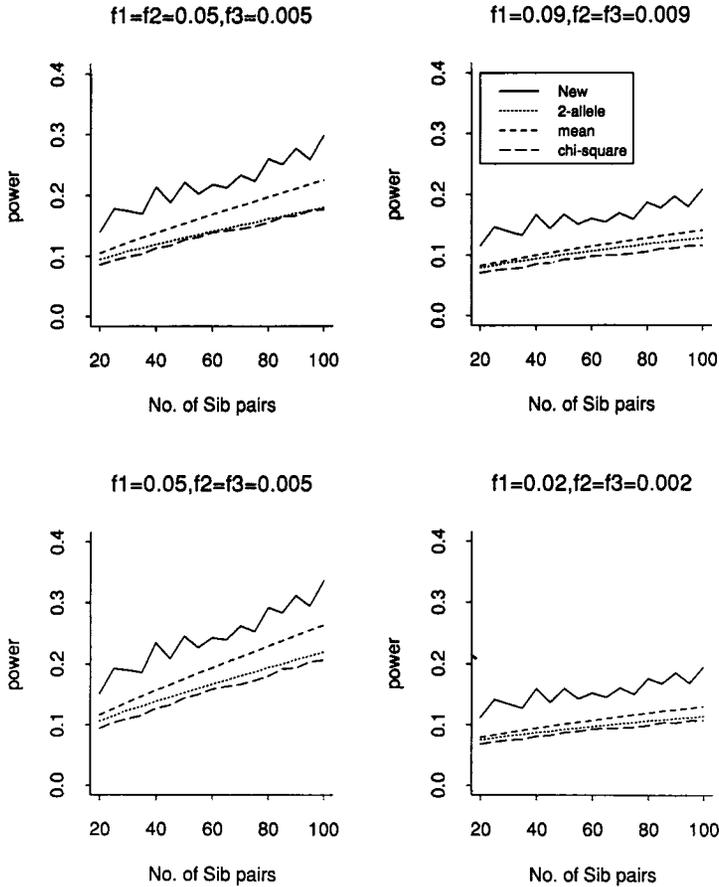


Fig. 3. Exact power for the new test (T), the two-allele test, (t_1), the mean test, (t_2), and the chi-squared goodness of fit test (Y) for a nominal 5% level of significance, complete linkage, $\theta = 0.25$, population frequency, $K = 0.01$.

The regions and values of the test statistic may be seen in Figure 1. T is continuous as $(\hat{p}_0, \hat{p}_1, \hat{p}_2)$ varies. The asymptotic distribution of T can be calculated exactly using the asymptotic multivariate normality of $(\hat{p}_0, \hat{p}_1, \hat{p}_2)$ but its form is not tractable. The exact distribution of T under the null hypothesis may be determined exactly for all values of n that are likely to occur in practice, because the computation is relatively inexpensive. A program is available from the author which can determine exact critical values for any desired size of test. Alternatively, Table I gives critical values based on $n = 150$.

POWER

We make the power comparisons as in Blackwelder and Elston [1985], setting $\theta = 0$, $K = 0.01$, and the significance level to 5%. Due to the discreteness of all the test statistics, a randomized decision rule was used to set the size at exactly 5% to

enable a fair comparison, although one should not do this in practice. The results are shown in Figure 2. The new test is more powerful than the others, the margin depending on the setting. The t_{max} statistic proposed by Schaid and Nick [1990] was close to the best of t_1 and t_2 so the new statistic outperforms it also. The lack of monotonicity in the power curve can be ascribed to the discrete and composite nature of the test statistic.

Other combinations of the parameters have been tried with the new test proving superior in every case. In Figure 3, we show the same power comparisons for $\theta = 0.25$. The difference in power between the new test and the older tests is more substantial.

DISCUSSION

We have shown that the new test statistic based on restricting the regular chi-squared goodness of fit statistic is more powerful for the detection of linkage than the previously considered tests.

It would also be possible to develop a test based on the likelihood ratio test, again restricting the alternative as above. This would be asymptotically equivalent to the chi-squared-based statistic, but its relative performance might vary in small samples.

If more assumptions can be made about the penetrances and population prevalences, this would result in a smaller feasible region for (p_0, p_1, p_2) and hence a more powerful test.

In cases of sibships with more than two members, the data consist of all possible pairs and hence mutual independence is lost. Asymptotically, this is of no concern as has been demonstrated by Blackwelder and Elston [1985], but there would be some impact on the exact critical values for small sample sizes [see Hodge, 1984].

REFERENCES

- Blackwelder WC, Elston RC (1985): A comparison of sib-pair linkage tests for disease susceptibility loci. *Genet Epidemiol* 2:85-97.
- Concannon P, Wright JA, Wright LG, Sylvester DR, Spielman RS (1990): T-cell receptor genes and insulin-dependent diabetes mellitus (IDDM): No evidence for linkage from affected sib pairs. *Am J Hum Genet* 47(1):45-52.
- Cox NJ, Baker L, Spielman RS (1988): Insulin-gene sharing in sib pairs with insulin-dependent diabetes mellitus: No evidence for linkage. *Am J Hum Genet* 42(1):167-172.
- Cudworth AG, Woodrow JC (1975): Evidence for HLA-linked genes in juvenile diabetes mellitus. *Br Med J* 3:133-135.
- Day NE, Simons MJ (1976): Disease susceptibility genes—Their identification by multiple family studies. *Tissue Antigens* 8:109-119.
- de Vries RRP, Fat RFMLA, Nijenhuis LE, Van Rood JJ (1976): HLA-linked genetic control of host response of *Mycobacterium leprae*. *Lancet* 2:1328-1330.
- Fimmers R, Neugebauer M, Dennert J, Wienker T, Bauer MP (1989): Association and sib-pair analysis for the HLA, Gm, Km, and insulin polymorphisms in multiplex IDDM families. *Genet Epidemiol* 6:107-112.
- Fishman PM, Suarez B, Hodge SE, Reich T (1978): A robust method for the detection of linkage in familial diseases. *Am J Hum Genet* 30:308-321.
- Fletcher R (1987): "Practical Methods of Optimization." 2nd Ed. New York: John Wiley & Sons.
- Forman D, Oliver RT, Brett AR, Marsh SG, Moses JH, Bodmer JG, Chilvers CE, Pike MC (1992): Familial testicular cancer: A report of the UK family register, estimation of risk and an HLA class I sib-pair analysis. *Br J Cancer* 65:255-262.

- Green JR, Woodrow JC (1977): Sibling method for detecting HLA-linked genes in disease. *Tissue Antigens* 9:31–35.
- Hodge SE (1984): The information contained in multiple sibling pairs. *Genet Epidemiol* 1:109–112.
- Holmans P (1993): Asymptotic properties of affected-sib-pair linkage analysis. *Am J Hum Genet* 52:362–374.
- Jeunemaitre X, Rigat B, Charru A, Houot AM, Soubrier F, Corvol P (1992): Sib-pair linkage analysis of renin gene haplotypes in human essential hypertension. *Hum Genet* 88(3):301–306.
- Knapp M (1991) A powerful test of sib-pair linkage for disease susceptibility. *Genet Epidemiol* 8:141–142.
- Lehmann EL (1986): "Testing Statistical Hypotheses." 2nd Ed. New York: John Wiley & Sons.
- Majumder P, Pal N (1987) Nonrandom segregation: Uniformly most powerful tests and related considerations. *Genet Epidemiol* 4:277–287.
- Schaid DJ, Nick TG (1990): Sib-pair linkage tests for disease susceptibility loci: Common tests vs. the asymptotically most powerful test. *Genet Epidemiol* 7:359–370.
- Schaid DJ, Nick TG (1991): A reply to "A powerful test of sib-pair linkage for disease susceptibility." *Genet Epidemiol* 8:142–143.
- Suarez BK, Reich T, Trost J (1976): Limits of the general two-allele single locus model with incomplete penetrance. *Ann Hum Genet* 40:231–244.
- Suarez BK, Rice J, Reich T (1978): The generalised sib-pair IBD distribution: Its use in the detection of linkage. *Ann Hum Genet* 42:87–94.
- Weitkamp LR, Stanser HC, Persad E, Flood C, Guttormsen S (1981) Depressive disorders and HLA: A gene on chromosome 6 that can effect behavior. *N Engl J Med* 305:1301–1306.

APPENDIX

To obtain the bound $3p_1/2 + p_2 \geq 1$ observe that

$$\frac{p_2 - 1/4}{p_1 - 1/2} = \frac{V_A + (\varphi + 1/2)V_D}{-2(\varphi - 1/2)V_D}.$$

The right-hand side will be maximized for $\theta = 0$, i.e., $\varphi = 1$, so an upper bound on the left-hand side is

$$-\frac{V_A}{V_D} - \frac{3}{2}.$$

Now V_A/V_D may be made arbitrarily small by setting $\theta = 0$, $p = \epsilon$, $f_1 = 1$, $f_2 = f_3 = \epsilon^2$

$$\frac{V_A}{V_D} = -2\frac{\epsilon}{\epsilon - 1}$$

which tends to zero as $\epsilon \rightarrow 0$. This gives us the bound $3p_1/2 + p_2 \geq 1$. To show that it is tight, set $\theta = 0$, $p = \epsilon$, $f_2 = f_3 = \epsilon^2$ where ϵ is small. Ignore terms of order higher than ϵ^2 to obtain (after some algebra)

$$p_1 = \frac{2\epsilon(2\epsilon f_1 + \epsilon - 2\epsilon^2 f_1 + f_1^2)}{\epsilon^2 f_1^2 + 6\epsilon^2 f_1 + 4\epsilon^2 + 2\epsilon f_1^2 + f_1^2}$$

and

$$p_1 = \frac{\epsilon^2 + f_1^2}{\epsilon^2 f_1^2 + 6\epsilon^2 f_1 + 4\epsilon^2 + 2\epsilon f_1^2 + f_1^2}.$$

Note that

$$3p_1/2 + p_2 = \frac{4\epsilon^2 + 6\epsilon^2 f_1 + f_1^2 + 3\epsilon f_1^2}{\epsilon^2 f_1^2 + 6\epsilon^2 f_1 + 4\epsilon^2 + 2\epsilon f_1^2 + f_1^2}$$

which tends $\downarrow 1$ as $\epsilon \rightarrow 0$. Now since

$$\lim_{\epsilon \rightarrow 0} p_1|_{f_1=1} = 0 \quad \lim_{\epsilon \rightarrow 0} p_1|_{f_1=0} = 1/2$$

and

$$\lim_{\epsilon \rightarrow 0} p_2|_{f_1=1} = 1 \quad \lim_{\epsilon \rightarrow 0} p_2|_{f_1=0} = 1/4.$$

It is possible to choose an f_1 such that (p_1, p_2) comes arbitrarily close to the line $3p_1/2 + p_2 = 1$ for a given p_1 between 0 and 1/2.

To show that the other two bounds are tight, similar arguments apply. For $p_1 \leq 1/2$, the boundary is described by setting $\theta = 0, p = \epsilon, f_1 = f_2 = 1$ and letting f_3 vary between 0 and 1 for ϵ small. For $p_1 + p_2 \leq 1$, the boundary described by setting $\theta = 0, p = \epsilon, f_1 = 1, f_3 = \epsilon^2$ and letting f_2 vary between 0 and 1.