PREFERENTIAL MATING IN SYMMETRIC MULTILOCUS SYSTEMS: LIMITS FOR MULTIALLELISM AND FOR MANY LOCI*

J. RAPER

Department of Genetics, Cambridge University, Cambridge CB2 3EH, England

AND

Department of Mathematics, Stanford University, Stanford, California 94305

Manuscript received August 13, 1981
Accepted October 31, 1981

ABSTRACT

Models in which general forms of preferential mating have been superimposed on the framework of the symmetric heterozygosity selection regime have been examined previously with respect to the existence and local stability of a central polymorphic equilibrium. The results are now extended to produce the limiting form of the stability conditions in two cases: First, where the number of alleles per locus is assumed to be very large; second, where the number of loci affecting the character is very large. It is argued that some type of frequency dependence in the mating pattern must be included, and a particular case is examined in detail. It is shown that multiallelism is ambiguous in its effect on stability, while an increasing number of loci, at least under zero linkage, leads to a simple stability condition which is analogous to the one-locus heterosis principle. Assortative mating appears to be more likely to produce a stable central polymorphism under high levels of allelism than is sexual selection, but is relatively very much weaker than sexual or viability selection if the number of loci involved is large.

IN a previous paper (KARLIN and RAPER, this issue) we have discussed models which attempt to consider nonrandom mating in the context of a trait genetically determined by several loci. This is achieved by considering preferential mating incorporated into one of the better studied classes of multilocus selection schemes, the symmetric heterozygosity models (KARLIN 1979; KARLIN and AVNI 1981): These, in turn, are a generalization of the symmetric models for two or three loci analyzed by several authors (KIMURA 1956; LEWONTIN and KOJIMA 1960; KARLIN and FELDMAN 1970, among others). A full description of the combined model is contained in the preceding paper. Briefly, we consider a large diploid population, sexually reproducing in discrete generations, where viability selection is acting on, and preferential mating is occurring with respect to, a character which is determined according to which of several loci are in a heterozygous state. The exact alleles involved do not matter, only whether or not those of each homologous pair are identical. Consequently, there is total symmetry be-

* Supported in part by NIH Grant 5R01 GM10462-18 and NSF Grant MCS79-26310.
between alleles at any locus, and as a result it is possible to establish the existence of a central equilibrium where every possible genotype is present at equal frequency. Conditions for local stability of this polymorphism, \( c^* \), were also obtained, and were seen to be the same as in the viability-only model (Karlin and Avni 1981), except for the need for a new "generalized fitness" \( (\gamma) \) to replace the plain viabilities. As noted there, the involvement in \( \gamma \) of the parameters \( m \) (the number of alleles per locus is assumed the same for each locus) and \( n \) (the number of loci), means that the consequences of having large values of \( m \) or \( n \) diverge from those of similar considerations in the viability-only case. It is to the analysis of these consequences that this paper is directed.

Many examples are known of polymorphisms in natural populations where several alleles are segregating (for instance, XdH, the HLA complex and other serological markers); this raises the question of whether high levels of allelism are generally beneficial towards the maintenance of polymorphism (see the simulation studies of Lewontin, Ginzburg and Tuljapurkar 1978). The approach here is to examine the limiting form for the stability conditions for \( c^* \) as \( m \to \infty \). On the other hand, a motivation for examining these multilocus models is that when the number of loci, \( n \), is very large, the situation parallels that of a quantitative trait, although the genetic basis for inheritance is still exhibited more clearly than in many polygenic models. It is, therefore, of interest to see what becomes of the local stability conditions as \( n \to \infty \). Two major problems arise: First, in both cases the population will become more homogeneous—as \( m \to \infty \) it is clear that most of the population will end up in the class of individuals heterozygous at every locus, and as \( n \to \infty \) the laws of large numbers ensure that near \( c^* \) the population clusters more and more closely around the average proportion of heterozygous loci. This leads to some increasingly rare classes and the need for some care in handling the question of rare male mating advantage. Second, in the case of increasing \( n \), the very means of phenotypic classification is subject to continual alteration so that some consistent way of defining the viability and mating parameters is required.

Mating advantage, and in particular that of the rare male, is one facet of frequency-dependence in the expression of mating preferences. If the population is divided into several phenotypic classes, \( C_0, C_1, \ldots, C_n \), then a preferential mating scheme may be described by a matrix \( A = (a_{ij}) \), where the entry \( a_{ij} \) gives the proportion of females in \( C_i \) who mate with a \( C_j \) male because of some selective preference (a formulation adapted from that in Karlin, 1978). This matrix may well depend on the frequencies of the various phenotypic classes (say, \( z_0, z_1, \ldots, z_n \)), the exact mode of dependence reflecting the supposed behavioral mechanism controlling the expression of preference. If \( v(j) \) is defined to be the proportion of matings undertaken by \( C_j \) males relative to their frequency in the population \((z_j)\), it can be seen that the contribution to \( v(j) \) from \( C_i \times C_j \) preferential matings is \( a_{ij} \frac{z_i}{z_j} \). Clearly, \( v(j) \), which is a measure of the mating success of \( C_j \) males, will become unbounded as \( z_j \to 0 \) unless \( a_{ij} \), through its dependence on the \( z \)'s, also tends to zero. (The only exception to this argument is when \( i = j \): thus,
assortative mating, where the only nonzero terms are on the diagonal, does not suffer from this possibility.) This contingency is biologically unrealistic since there must be at some stage a limit to the number of fertilizations that any one male can perform. Very high mating advantages have been observed, and it is certainly conceivable that in a finite population one male could be responsible for all fertilizations, but since this is a model of an effectively infinite population some bound on the mating success coefficient must be assumed.

It was remarked above that both of the limiting cases to be studied here involve a number of increasingly rare phenotypic classes, so that these arguments imply the need for some frequency-dependence to be included. This is reinforced by looking at matters from the female's point of view: there must be some limit to how long she will search for a virtually nonexistent type of male, however desirable that type may be. (This point shows why frequency-dependence should also be included in pure assortative mating models although the problem of unbounded mating success does not arise there.) The precise functional dependence of the matrix, $A$, on the frequencies, $z_0, \ldots, z_n$, may be deduced from the presumed behavioral aspects in the selection of mates. For instance, amongst others which have been proposed, there are the encounter models of O'Donald (1978); see also Karlin and Raper (1979); Raper, Karlin and O'Donald (1979); the searching-time models of Eshel (1979), and the "preference function" models of Wagener (1976). Indeed, there is one very simple form which is a special case common to all three of these types of model, namely that in which the proportion of preferential matings is directly proportional to the frequency of the type of male preferred: $a_{ij} = b_{ij}z_j$, where $b_{ij}$ are constants. In this context, the frequency of $C_i \times C_j$ preferential matings (i.e., ignoring those occurring at random with unselective females) is $b_{ij}z_i z_j$, so that $b_{ij}$ can be interpreted as the probability of a mating between a $C_i$ female and a $C_j$ male conditional on their encountering one another. In view of its mathematical simplicity and its plausible probabilistic interpretation, this is the form of frequency-dependence employed below.

The second problem is approached by a device paralleling that used by Karlin (1979) (see also Karlin and Avni 1981) for the viability-only version of the symmetric heterozygosity model. The aim is to see what becomes of the conditions for stability of $c^*$ as the number of loci becomes larger, which means comparing the results from a sequence of models with increasing values of $n$; but the same parameter specifications cannot be used throughout since they are defined on the criterion of heterozygous loci, and the number of available loci is changing. Some underlying, unifying factor must be found to allow direct comparison of results from models with different values of $n$. A simple and natural answer can be found in the case of the completely symmetric heterozygosity model where the phenotypic classes (and, therefore, the parameter values) are independent of the positions of heterozygous loci and are determined merely by how many there are: this is to regard the basic criterion as being the proportion of loci which are in heterozygous state and postulating some underlying function relating this proportion to the required parameter. For example, to define the
relative viability in females, assume a viability function \( \phi \) so that for \( C_k \), the phenotypic class which has \( k \) out of \( n \) loci heterozygous, the viability is given by \( \phi(k/n) \). This is usually adequate since it is only the overall shape of \( \phi \) which is of interest. For instance, with a polygenic trait the number of loci involved is unlikely to be known, but certain qualitative features, such as heterozygote vigor or epistatic effects, can be reflected by taking \( \phi \) to be increasing or concave, etc. In this case, there is no ambiguity: since the viabilities are presumed frequency-independent there is no choice but to take the unifying function \( \phi \) to be the viability itself. However, in defining the mating parameters, with their inherent frequency-dependent character, the question arises as to what form of parameterization should be linked to a fixed, common function. The existence of very rare phenotypic classes argues that a necessarily frequency-dependent formulation should be chosen, and its simple probabilistic interpretation again recommends that \( b_{ij} \) should be connected to an underlying mating function. This is developed in Section 2.

The question of a high level of allelism \( (m \to \infty) \) is dealt with in Section 1, where it is also convenient to quote the results which will be needed from the first paper (KARLIN and RAPER 1982). For the sake of conformity with Section 2, only the completely symmetric model is considered here, so that phenotypic class is determined solely by the degree of heterozygosity. Section 2 is devoted to the derivation of a limiting condition for the stability of \( c^* \) as \( n \to \infty \), but only in the single case of free recombination between all loci. These results are then discussed and compared.

1. **High Allelism**: A succinct description of the stability conditions for the central polymorphism, \( c^* \), relies on the definition for each phenotypic class \( C_k \) of a single quantity combining the viabilities of either sex with the mating scheme to give a generalized "fitness", \( \gamma(k) \). Recalling that \( s_k \) and \( t_k \) give the female and male relative viabilities, respectively, for \( C_k \), and that \( \sigma \) and \( \tau \) represent their respective population mean values, then

\[
\gamma(k) = \frac{1}{2} \left[ \frac{s_k}{\sigma} + \frac{t_k}{\tau} (1-A.) + \frac{a_{jk}}{z_k^*} \right].
\]

There \( z_k^* \) denotes the equilibrium frequency of the class \( C_k \), and the dotted subscript stands for the average with respect to adult female frequencies.

Since we will now be working with the frequency-dependent form of the matrix \( A \) specified in the Introduction, it is now convenient to rewrite \( \gamma(k) \) directly in terms of the quantities \( b_{jk} \). Note, however, that \( z_k^* \) is the frequency of \( C_k \) in the juvenile population, while it is reasonable to suppose that \( a_{jk} \) will in fact be proportional to male adults' frequencies, so it is more natural to use the designation

\[
a_{jk} = b_{jk} t_k \tau^{-1} z_k^*.
\]

Then \( a_k = \sum s_j z_j^* \sigma^{-1} a_{jk} = t_k \tau^{-1} z_k^* b_{jk} \) and \( A. = \sum a_k = b. \) (where the second dotted
subscript refers to averaging over the male adult population. Substituting in (1.1) gives

$$\gamma(k) = \frac{1}{2} \left[ \frac{s_k}{\sigma} + \frac{t_k}{\tau} (1-b_{..} + b_k) \right].$$

(1.3)

It was remarked in Karlin and Avni (1981) that linear rescaling of the $\gamma$'s does not affect the conditions for local stability of $c^*$ (although adding a constant to all the $\gamma$'s will reduce the strength of stability). For the purpose of deciding stability, it can be seen from (1.3) that, in the absence of any viability selection, the $\gamma(k)$'s may be replaced by $b_{..k}$: for in that case $\gamma(k) = \frac{1}{2} [2 - b_{..} + b_k]$ which is linearly dependent on $b_{..k}$.

Consider, for example, the standard form of sexual selection where preferences are unaffected by the females’ own phenotypes: the mating scheme is, therefore, described by a matrix where each row is the same, say $b_{jk} = c_k$. The $b_{..k} \sigma^{-1}(\Sigma_j \gamma_j^* b_{jk} = c_k$. In view of the statement just made above, we see that when this type of frequency-dependent sexual selection acts on its own the stability of the central polymorphism can be decided by replacing the fitness parameters $\gamma(k)$ by the sexual selection parameters $c_k$. This contrasts with the nonfrequency-dependent mating schemes where it is assortative mating which mimics viability selection, but in either case the similarity is coincidental being contingent on the particular choice of dependence function.

We are now in a position to examine, for a general preferential mating scheme, the consequence of high levels of allelism. Consider first free recombination: replacing $\gamma(k)$ by $b_{..k}$ in the condition given by Karlin and Avni, $c^*$ is stable if $\Sigma_k (m-1)^k (\Sigma_j^{n-1}) [b_{..k} - b_{..k+1}] < 0$. At $c^*$ the frequency of the $j^{th}$ phenotypic class, $C_j$, is seen to be $z_j^* = m^{-n} (m-1)^j (\gamma_j^*), so performing the average implied by the dotted subscript the condition is

$$\Sigma_{k,j} (m-1)^{k+j} (\Sigma_j^{n-1}) (\gamma_j^* [b_{jk} - b_{..j,k+1}] < 0. \tag{1.4}$$

For large $m$ this condition will be dominated by the highest power of $(m-1)$, which is attained by taking $k = n-1$ and $j = n$: this has coefficient $b_{n,n-1} - b_{nn}$. So for large $m$ the stability condition is

$$b_{n,n-1} < b_{nn}. \tag{1.5}$$

The opposite extreme is given by a trait controlled by loci between which there is no recombination. Substituting $b_{..k}$ for $\gamma(k)$ in the zero recombination conditions shows that $c^*$ is stable if

$$\Sigma_{i,k,l} (-1)^i (m-1)^k l^j l^+ (\gamma_j^* (\Sigma_j^{n-1}) (\gamma_j^* b_{ik} < 0 \text{ for } 1 \leq l \leq n. \tag{1.6}$$

In the $l^{th}$ condition, the highest power of $(m-1)$ is $(2n-l)$ which is obtainable by $i = n, j \leq l$ and $k = n-l+j$. The limiting conditions for large $m$ are, therefore,
It is immediately visible that the conditions under which preferential mating alone can maintain a stable central polymorphism when there are very many alleles at each locus are considerably more restrictive when there is no recombination \((1.8)\) than for free recombination \((1.5)\). This reflects the general property of the symmetric heterozygosity models that greater recombination facilitates the stability of \(c^*\).

There is also a sharp contrast in the outcomes from the two special mating schemes often studied in the literature, assortative mating and sexual selection. As mentioned above, sexual selection is described by matrix of the form \(b_{jk} = c_k\): \((1.5)\) is then simply \(c_{n-1} < c_n\), so that stability under free recombination is quite possible but not definite. \((1.8)\) produces conditions on the \(c_k\)'s which show that stability of \(c^*\) (as \(m \to \infty\)) for zero recombination is possible, but considerably less likely. On the other hand, assortative mating entails a scheme where the only preferences are between like phenotypes, yielding a diagonal matrix: thus, \(b_{aa} = b_{n1} = \ldots = b_{n,n-1} = 0\), and \(b_{nn} > 0\) (in the generic case). So \((1.5)\) is necessarily satisfied, while \((1.8)\) must fail \((l = 2\) never works).

Finally, it can be noted how, at both extremes of recombination, the stability conditions under high levels of allelism involve only the last row of the mating matrix, that is, the preferences of the females with the highest levels of heterozygosity. This is to be expected, for the presence of many alleles at each locus will make the state of heterozygosity more common, so that females in the phenotypic class of greatest heterozygosity will tend to dominate the population and it is their sexual preferences which will control the stability outcome. Indeed, for free recombination the only condition in the limit is that \(c_n\) females should prefer \(c_{n-1}\) males more than the next most common class, \(C_{n-2}\).

2. Many loci: The aim of this section is to discover what form the conditions for local stability of \(c^*\) assume when \(n\) is large. We discussed in the INTRODUCTION a reasonable means of coping with the problem of comparing models with different numbers of loci in spite of the resultant change in phenotypic classification. This is now set down in detail.

Let \(\phi\) and \(\psi\) be the female and male viability functions, respectively. Thus, for any given number of loci \(n\), the relative viabilities of \(C_k\) individuals (i.e., those in which \(k\) out of \(n\) are heterozygous) are respectively,

\[
s_k = \phi(k/n) \quad \text{and} \quad t_k = \psi(k/n) .
\]

These are positive functions defined on the unit interval \([0,1]\), and will be assumed to be smooth and bounded above, and away from zero below.

Let \(g(x,y)\) be the conditional mating function. Again, \(g\) is assumed to be a smooth function on the unit square \((0 \leq x,y \leq 1)\), and, in view of its probabilistic interpretation, to take values between 0 and 1. \(g(x,y)\) is to be the conditional probability of a successful mating between a selectively mating female and a
male who have respectively a fraction $x$ and $y$ of their loci heterozygous. In previous terminology this means

$$b_{jk} = g(j/n, k/n) \quad (2.2)$$

The procedure is then to insert these parameter specifications (2.1) and (2.2) into the stability conditions, and to examine the limit as $n \to \infty$. Note that we only attempt this here in the particular case of totally free recombination. To review the notation and results to be used: $m$ is the number of alleles per locus, taken to be the same for each locus; $\sigma(\tau)$ is the average value of $s_k (t_k)$ when the population is at the equilibrium, $c^*$ (when using $\sigma$'s from models with different $n$'s we may write $\sigma_n, \text{etc.}$); $M$ is the total number of possible gametes ($M = m^n$); $z^*_k$ is the equilibrium frequency of the phenotypic class $C_k$, which is $M^{-1} (m-1)^k \binom{n}{k}$; and the (single) stability condition under free recombination is that

$$w = \sum_k (m-1)^k \binom{n-1}{k} [\gamma(k) - \gamma(k+1)] < 0 \quad . \quad (2.3)$$

Indeed, the largest eigenvalue of the local stability linear approximation is $\lambda = 1 + \frac{1}{M} w$. Use will be made of the following standard result.

**Lemma 1.** Let $f$ be a continuous and bounded real function, and let $X_n$ be a sequence of random variables converging to the value $p$ in probability, then $E[f(X_n)] \to f(p)$.

This is useful here since the binomial coefficients is $z^*_k$ and $w$ can be interpreted as probabilities of suitable binomial random variables.

**Lemma 2.** Let $f$ be a continuous function on the unit interval $[0,1]$, let $0 \leq p \leq 1$, and let $p_k = \phi^k (1-p)^{n-k} \binom{n}{k}$. Then $\sum_k p_k f(k/n) \to f(p)$ as $n \to \infty$. (See appendix for an outline of the proof.)

If we now take $p = 1 - 1/m$, then it is easy to see that $z^*_k = p_k$. Recalling the definition of $\sigma_n$ and the specification of the $s_k$'s (2.1),

$$\sigma_n = \Sigma s_k z^*_k = \sum_{k=0}^n p_k \phi(k/n) .$$

The previous lemma then implies that $\sigma_n \to \phi(p)$ (and similarly, $\tau_n \to \psi(p)$). This reflects the observation made earlier that when the number of loci is large the "laws of averages" ensure that most of the population have the average level of heterozygosity. This will be $p$ since, given one allele, the chances that the other allele at that locus is different is $\frac{m-1}{m} = p$; so the mean viability converges to the actual viability at $p$.

To illustrate the behavior of $w$ for larger $n$, we have:

**Lemma 3.** Let $f$ be a continuously differentiable function on $[0,1]$, let $w$ be as in (2.3), and let $p = \frac{m-1}{m}$. If $\gamma(k) = f(k/n)$, then

$$\frac{n}{M} w \to - (1-p)f'(p) .$$

(See appendix for proof.)
This description, $\gamma(k) = f(k/n)$, is precisely what is needed to handle the viability part of $\gamma$ [compare with (2.1)]. The stability discriminant, $w$, is linear in $\gamma$ so the components may be dealt with separately. For instance, reference to (1.3) shows that part of the influence of the preferential mating scheme comes through the term $b_{jk}$. In the new notation, this is

$$b_{jk} = \Sigma_i \sigma_i s_i z_i^* b_{ijk} = \sigma_i \Sigma_j p_j s_j b_{ijk} = \sigma_i \Sigma_j p_j \phi \left( \frac{i}{n} \right) g \left( \frac{j}{n}, \frac{k}{n} \right).$$

How this affects the behavior of $w$ as $n \to \infty$ can be examined in much the same way as before with some minor adaptations for the change to two dimensions. All these influences can then be combined to give the overall stability condition:

**Theorem.** Given viability functions $\phi$ and $\psi$ (2.1) and mating function $g$ (2.2), as $n \to \infty$ the dominant eigenvalue governing the stability of $c^*$ under free recombination is asymptotically

$$\lambda \approx 1 - \frac{1}{2mn} \left[ \frac{\phi'(p)}{\phi(p)} + \frac{\psi'(p)}{\psi(p)} + \frac{\partial g}{\partial y}(p,p) \right],$$

and the limiting condition for local stability is

$$\frac{\phi'(p)}{\phi(p)} + \frac{\psi'(p)}{\psi(p)} + \frac{\partial g}{\partial y}(p,p) > 0. \quad (2.5)$$

**DISCUSSION**

The previous section has shown the derivation of a condition for the local stability of the central polymorphism in a multilocus model with an arbitrarily large number of loci involved. The condition is simple in form and easy to interpret despite the allowance of general viability and mating functions. The development has relied solely on direct application of Mendelian segregation principles: the main structural assumptions are the means of determination of the phenotype and the absence of any linkage of loci.

Preliminary examination of the concluding theorem of Section 2 shows two immediate consequences of the greater homogeneity of the population when $n$ is large: First, the condition for stability comes to depend only on the behavior of the three functions ($\phi$, $\psi$, $g$) near point $p$; second, the strength of stability, as measured by how far the dominant eigenvalue is below 1, steadily decreases. A similar phenomenon was noted in Section 1 as $m \to \infty$: high levels of allelism imply that heterozygosity is the norm, and the stability weakens and comes to depend only on the preferences of full heterozygotes. The weakening of stability has the effect of making the conclusions for large $n$ less robust against structural perturbations, but the rate at which $\lambda$ approaches unity is only algebraic so that qualitative inferences may still be valid for relatively big values of $n$.

The essence of Lemma 3 is that $c^*$ will be locally stable if the generalized fitness $\gamma$ is increasing near $p$ (or, strictly speaking, near $k = np$). As mentioned in the previous paper, this fitness takes into account both the viability rates and the mating success of males. In these terms, it can then be said that there is a locally stable central polymorphism if fitness increases with heterozygosity, at
least near the population's most common levels. It is clear from the condition (2.3) that increasing $\gamma$ is sufficient but not necessary for stability of $c^*$. The conclusion of this analysis is that for large $n$ the condition narrows to the point where it is necessary and sufficient for $\gamma$ to be increasing at $p$. We therefore have an analog to the one-locus, two-allele heterosis principle that heterozygote advantage maintains a polymorphism.

The expression (2.5) shows that the derivative of $\gamma$ can be viewed as the sum of three component derivatives. The first two are the rates of change with respect to the heterozygosity proportion of the viabilities in the two sexes. These are appropriately normalized by the mean population values; since there is no inter-sex competition they are normalized separately. The third term may be seen to be the rate of change of $\nu$ (the coefficient of mating success defined in the INTRODUCTION. The three components—female and male relative viability, and mating advantage—contribute equally. Note that for the mating pattern given by $g$, it is only the marginal behavior with the female coordinate fixed at $x = p$ which is relevant.

Thus, pure sexual selection, where preferences are independent of the females' own type, would be represented by a mating function which does not involve $x$: $g(x,y) = g(y)$. As remarked in Section 1, sexual selection then behaves very much like viability selection with regard to the stability of $c^*$. Pure assortative mating, in the sense that only matings between individuals of exactly the same phenotype have increase likelihood, does not have a realistic meaning as $n$ becomes large since the subdivision of the population by phenotype becomes correspondingly finer. The ability to select for only one phenotypic value in a preferred mate then becomes rapidly less reasonable. The principle, however, of more frequent matings between individuals of similar type is incorporated by a mating function with the qualitative characteristic that $g(x,y)$ should be greatest along the diagonal $x = y$, tailing off on either side. In particular, $g(p,y)$ would usually have its maximum value at $y = p$, so that the partial derivative at $(p,p)$ in (2.5) is zero. More detailed analysis of the convergence in Lemma 3 shows that the limiting involvement of $g$ in $w$ is then through its second and third partial derivatives, but this effect is an order of $n$ smaller than that of the viability selection: so assortative mating would have relatively little influence on the stability of $c^*$ if the number of loci is large.

We have seen that, at least in the case of loose linkage, examination of the limits as $m$ or $n$ gets larger allows great simplification of the stability conditions and hence a clearer picture of the major influences on the possibility of maintaining a central polymorphism in these extreme cases.

This paper describes part of my research carried out at Stanford University under the supervision of Professor S. Karlin towards the requirements for a Ph.D.

LITERATURE CITED


APPENDIX

Proof of Lemma 2. Let $Y_1, Y_2, \ldots$ be a sequence of independent, identically distributed random variables, with \(\text{Prob}(Y_1 = 1) = p, \text{Prob}(Y_1 = 0) = 1-p\). Let $X_n = \frac{1}{n} \sum_1^n Y_j$; then the Weak Law of Large Numbers implies that $X_n \to p$ in probability. Also, \(\text{Prob}(X_n = k/n) = \text{Prob}(\sum_1^n Y_j = k) = p_k\). Therefore, $\sum p_k f(k/n) = E[f(X_n)]$, and the result follows from the first lemma.

Proof of Lemma 3. Since \(\binom{n-1}{k} = \binom{n}{k} - \binom{n}{k+1}\), we have $\frac{n}{M} w = n \sum_{k=1}^{n-1} p_k \left(1 - \frac{k}{n}\right)$.

Now, by hypothesis, $f'$ is uniformly continuous, and also as $n \to \infty$, $\left|\frac{k}{n} - \xi_{k,n}\right| < \frac{1}{n} \to 0$ so the second sum above approaches zero. Applying the preceding lemma to the function $(1-x)f'(x)$, the result follows.