

Einige theoretische Ergebnisse zur pflanzenzüchterischen positiven Massenauslese bei Berücksichtigung von Konkurrenzeffekten

M. Hühn

Institut für Pflanzenbau und Pflanzenzüchtung der Christian-Albrechts-Universität Kiel (BRD)

Some Theoretical Results on Positive Mass Selection in Plant Breeding with Special Consideration of Competitive Effects

Summary. To quantitatively investigate the competitive effects at the borders of neighbouring plots (using field experiments arranged in plots) an extensive series of field experiments, where the treatments were certain crosses of spruce, had been analysed by two different methods: 1) An analysis with almost complete elimination of the competitive effects at the borders of the plots; and 2) an analysis without explicit consideration and elimination of these competition-border-effects (see: Hühn 1974).

These studies on estimating competitive effects in plant stands quantitatively finally resulted in relations (inequalities) between the phenotypic total variances of these two methods of analysis, with (V^*) and without (V) eliminating the competition-border-effects: $V^* \geq V$. Furthermore, relations between different variance components were obtained from the analysis of variance computations of these two methods. The main purpose of the present paper is to draw some interesting theoretical conclusions, using only the inequality $V^* \geq V$, about 1) the covariance between treatment(genetic)-effects and competitive effects and 2) about the coefficients of determination and correlation-coefficients between the phenotypic values and the treatment-effects and between the phenotypic values and the competitive effects. To describe these relations quantitatively the ratio f of the competition variance to the variance of the treatment-effects is especially suitable.

In the special case $V^* = V$, which, however, has exceptional practical relevance, one obtains - for a large part of the possible interval ($0 \leq f \leq 2$) - for the ratio H of the variance of the treatment-effects to the total phenotypic variance (under certain assumptions this is the broad sense heritability), an interesting interpretation as coefficient of multiple determination R' (phenotypic value dependent on treatment-effects and competitive-effects). For $0 \leq f \leq 2$ this R' is explicitly totally independent of the magnitude of the competition variance.

In the main part of this publication, applications for breeding (positive mass selection) are discussed for the case $V^* = V$, which is of special practical relevance. Starting with a simple formula for the correlation-coefficient between the phenotypic values and the treatment effects, quantitative conditions for a possible application of positive mass selection in breeding are derived; with this the demand of a certain minimal value c of the corresponding coefficient of determination, which must be exceeded, is used as a condition for selection.

In the formulation of these conditions one obtains results about: 1) minimal values for H (dependent on c and on the number N of individuals, which must be selected); 2) derivation of f -intervals, where positive mass selection should not be applied (dependent on N , H and c , where an interesting special case arises for very large numbers of individuals ($N \rightarrow \infty$)); 3) the computation of the necessary minimal number of individuals which must be selected (dependent on f , H and c), where the case: "correlation between phenotypic values and treatment-effects significantly different from zero", which is of special interest from the point of view of practical breeding, is discussed in detail.

Zusammenfassung. Zur Untersuchung der Auswirkung der Parzellenränder-Konkurrenzwirkungen zwischen benachbarten Parzellen in parzellenweise angelegten Feldversuchen wurde eine umfangreiche Feldversuchsserie mit Fichtenkreuzungen verschiedenen Auswertungen unterzogen: 1) Eine Auswertung mit weitgehender Ausschaltung der Konkurrenzeffekte an den Parzellenrändern sowie 2) eine Auswertung ohne explizite Berücksichtigung und Eliminierung der Parzellenränder-Konkurrenzwirkungen (siehe Hühn 1974). Diese Untersuchungen zur quantitativen Einschätzung von Konkurrenzeffekten in Pflanzenbeständen führten schließlich - unter anderem - zu Größenbeziehungen (Ungleichungen) zwischen den phänotypischen Gesamtvarianzen dieser beiden Auswertungen mit (V^*) und ohne (V) Berücksichtigung der Parzellenränder-Konkurrenzwirkungen: $V^* \geq V$ sowie zu Größenbeziehungen zwischen verschiedenen Varianzkomponenten, die aus den varianzanalytischen Verrechnungen dieser beiden Auswertungen erhalten wurden.

In der vorliegenden Arbeit werden nun unter ausschließlicher Verwendung der Ungleichung $V^* \geq V$ interessante theoretische Folgerungen über 1) die Kovarianz zwischen Versuchsglied- und Konkurrenzeffekten und 2) über die Bestimmtheitsmaße bzw. Korrelationskoeffizienten zwischen phänotypischem Wert und Versuchsgliedeffekt sowie zwischen phänotypischem Wert und Konkurrenzeffekt abgeleitet.

Zur quantitativen Beschreibung dieser Zusammenhänge erweist sich das Verhältnis f der Konkurrenzvarianz zur Varianz der Versuchsgliedeffekte als besonders geeignet.

Im Sonderfall $V^* = V$, der jedoch von einer außerordentlichen praktischen Bedeutung ist, ergibt sich dabei für einen großen Teil des in Frage kommenden Bereiches ($0 \leq f \leq 2$) für den Anteil H der Varianz der Versuchsgliedeffekte an der phänotypischen Gesamtvarianz (unter bestimmten Voraussetzungen ist dies also die

'Heritabilität im weiteren Sinn') eine interessante Deutung als multiple Bestimmtheit R' (phänotypischer Wert in Abhängigkeit von Versuchsgliedeffekt und Konkurrenzeffekt), wobei R' für diesen Bereich explizit völlig unabhängig von der Größe der Konkurrenzvarianz ist.

Im Hauptteil der Arbeit werden dann für diesen praktisch äußerst bedeutsamen Fall $V^* = V$ züchterische Anwendungen (positive Massenauslese) diskutiert. Ausgehend von einer einfachen Formel für den Korrelationskoeffizienten zwischen phänotypischem Wert und Versuchsgliedeffekt (wobei die Forderung eines bestimmten zu überschreitenden Mindestwertes c des entsprechenden Bestimmtheitsmaßes als Selektionsbedingung benutzt wird) lassen sich Bedingungen für eine sinnvolle züchterische Anwendung der positiven Massenauslese ableiten und durch quantitative Beziehungen präzisieren.

Bei der Formulierung dieser Bedingungen kommt man zu Aussagen über: Mindestwerte für H (in Abhängigkeit von der Anzahl N der zu selektierenden Individuen und c), Angabe von für die positive Massenauslese nicht zulässigen f -Intervallen (in Abhängigkeit von N , H und c , wobei sich ein interessanter Sonderfall für sehr hohe Individuenanzahlen ($N \rightarrow \infty$) ergibt) und schließlich die Berechnung der mindestens notwendigen Anzahlen der zu selektierenden Individuen (in Abhängigkeit von f , H und c), wobei hier besonders der unter praktischen züchterischen Gesichtspunkten interessierende Fall: 'Signifikant von Null verschiedene Korrelation zwischen phänotypischem Wert und Versuchsgliedeffekt' diskutiert wird.

1. Einleitung

In einer früheren Veröffentlichung (Hühn 1974) wurde über experimentelle Konkurrenzuntersuchungen berichtet, die anhand einer Feldversuchsserie mit Fichtenkreuzungen des Instituts für Forstgenetik und Forstpflanzenzüchtung in Schmalenbeck (BRD) durchgeführt worden sind. Es handelte sich dabei um einen im Frühjahr 1959 an sechs verschiedenen Standorten der Bundesrepublik Deutschland mit je zwanzig zu prüfenden "Sorten" = Versuchsgliedern angelegten Versuch, der bei den in den Jahren 1959, 1963, 1966, 1969 und 1971 durchgeführten routinemäßigen Messungen des Merkmals 'Pflanzenhöhe' an allen sechs Standorten durch sehr große Konkurrenzwirkungen an den Parzellenrändern auffiel.

Der Versuchsplan war an allen Standorten ein Rechteckgitter mit drei Wiederholungen (also ein Dreisatz-Rechteckgitter ohne Wiederholung des Grundplans) mit fünf Blöcken à vier Versuchsgliedern je Wiederholung und mit $6 \times 6 = 36$ Pflanzen je (quadratischer) Parzelle.

Die zwanzig in den Versuchen verwendeten Versuchsglieder bestanden teils aus durch freies Abblühen erhaltenen Einzelstammabsaaten bestimmter ausgewählter Einzelbäume der Forstämter Westerhof (Harz) und Sundmo (Nordschweden), teils bestanden sie aus einer Reihe von Kreuzungen zwischen verschiedenen Sundmo-Pflanzen und einzelnen Westerhof-Pflanzen bzw. Einzelpflanzen des Schmalenbecker Institutes.

Auf nähere Angaben sowohl bezüglich dieses verwendeten Materials als auch der einzelnen Versuche soll hier verzichtet werden, da sie für die folgenden theoretischen Untersuchungen nicht benötigt werden (siehe hierzu Hühn 1974).

Um zu quantitativen Aussagen über die Größe der konkurrenzbedingten Parzellenrandwirkungen und den daraus resultierenden Verzerrungen zu kommen, wurden in Hühn (1974) jeweils zwei verschiedene Auswertungen der einzelnen Versuche vorgenommen: Zunächst die sogenannte 'Gesamtparzellenauswertung', bei der alle Pflanzen einer Parzelle, d.h. alle $6 \times 6 = 36$ Einzelpflanzen, benutzt wurden, und dann schließlich eine zweite Auswertung, die sogenannte 'Parzellenkernauswertung', bei der sämtliche Randpflanzen jeder Parzelle weggelassen wurden und wo nur der verbleibende Parzellenkern von $4 \times 4 = 16$ Einzelpflanzen in die Auswertung genommen wurde.

Aus unterschiedlichen Ergebnissen dieser beiden Auswertungen, z.B. bei den Versuchsglieder-Rangfolgen, bei den Mittelwertvergleichen anhand statistischer Signifikanztests, bei den Schätzwerten der verschiedenen Varianzkomponenten usw., lassen sich dann Schlüsse auf die quantitativen Beziehungen und verzerrenden Effekte dieser Parzellenrand-Konkurrenzwirkungen ziehen (siehe nähere Angaben hierzu bei Hühn 1974).

"Wie die numerischen Ergebnisse dieser Berechnungen zeigen, gilt für alle Versuchsorte und alle Untersuchungsjahre (bis auf zwei Ausnahmen, bei denen die beiden Varianzen jedoch praktisch als gleich angesehen werden können), daß die Gesamtvarianz bei der 'Parzellenkernauswertung' (V^*), d.h. bei weitgehender Ausschaltung der Konkurrenzeffekte an den Parzellenrändern, größer ist als die Gesamtvarianz bei der 'Gesamtparzellenauswertung' (V), bei der die Konkurrenzwirkungen an den Parzellenrändern weder ausgeschaltet noch explizit berücksichtigt werden. Unter Berücksichtigung und Einbeziehung der

Konkurrenzeffekte an den Rändern benachbarter Parzellen erhält man also eine kleinere Gesamtvarianz, d.h. die Konkurrenzwirkungen führen zu einer Nivelierung der Gesamtvariabilität" (Hühn 1974).

Aus den experimentellen Untersuchungen folgt also das Ergebnis:

$$V^* \geq V. \quad (1)$$

Dieses ist ein einigermaßen überraschendes Resultat, denn in vielen experimentellen Konkurrenzuntersuchungen resultierte die Auswirkung der Konkurrenz auf die Komponenten im Mischbestand in einer Erhöhung der Leistungsdifferenzierung gegenüber dem reinen Bestand. Dabei war es in diesen Untersuchungen meist so, daß die wüchsigen Typen profitierten und die weniger wüchsigen unterdrückt wurden, wodurch eine Vergrößerung der Unterschiede zustandekommt (Singh 1967, Geidel und Haufe 1968, Stern 1969). In unserem Fall führten jedoch die Parzellenrand-Konkurrenzwirkungen zu einer Verminderung der Gesamtvariabilität, woraus sich im folgenden einige interessante theoretische Konsequenzen ableiten lassen.

Hierbei werden zunächst Folgerungen aus dieser Beziehung (1) untersucht, während dann schließlich weitergehendere theoretische Ergebnisse sowie züchterische Anwendungen (positive Massenauslese) für den Spezialfall $V^* = V$ diskutiert werden. Ein entscheidendes Argument für die Behandlung dieses Sonderfalles liegt darin, daß dieser Fall $V^* = V$ - zumindest aufgrund der anfangs zitierten experimentellen Konkurrenzuntersuchungen - eigentlich überhaupt kein Spezialfall ist. Denn in den 31 hierbei verwendeten Versuchsauswertungen gilt zwar in 29 Fällen die Beziehung $V^* > V$, jedoch sind die Varianzenunterschiede so geringfügig, daß sie sich in keinem einzigen Fall als statistisch signifikant erweisen; auch in den beiden restlichen Fällen, in denen $V^* < V$ gilt, sind die Unterschiede nicht signifikant.

Die Ableitung von Folgerungen und züchterischen Anwendungen für den Fall $V^* = V$ kann nach diesen Feststellungen also wohl kaum als die Beschäftigung mit einem sehr speziellen und daher relativ uninteressanten Sonderfall angesehen werden.

Unter den in Hühn (1974) gefundenen und dort eingehend diskutierten Resultaten dieser experimentellen Konkurrenzuntersuchungen sind im Rahmen der folgenden Überlegungen - neben der Relation (1) zwi-

schen den Gesamtvarianzen bei beiden Auswertungen - zwei weitere Ergebnisse von Interesse:

"Der Variabilitätsanteil für 'Versuchsfehler' nimmt bei Übergang von der 'Gesamtparzellenauswertung' zur 'Parzellenkernauswertung' in allen Jahren und an allen Orten (bis auf eine einzige wahrscheinlich zufallsbedingte Ausnahme) in einem ganz extremen Maße zu" (Hühn 1974). (2)

Der Variabilitätsanteil für 'Versuchsglieder' ist bei der 'Parzellenkernauswertung' durchweg zum Teil wesentlich kleiner als bei der 'Gesamtparzellenauswertung'. Das heißt: "Bei der in der Praxis allgemein üblichen 'Gesamtparzellenauswertung', bei der die Konkurrenzwirkungen an den Parzellenrändern weder ausgeschaltet noch explizit berücksichtigt werden, wird der Variabilitätsanteil für 'Versuchsglieder' zum Teil weit überschätzt. Die Konkurrenz an den Parzellenrändern führt also zu einer stärkeren Differenzierung der Versuchsglieder; Versuchsgliederunterschiede werden stärker ausgeprägt und als Folge hiervon umfaßt die im Experiment berechnete Versuchsgliedervarianz eine konkurrenzbedingte Komponente. Die Nichtberücksichtigung der Konkurrenz an den Parzellenrändern führt also zu einer verzerrenden Verminderung der Versuchsfehlervariabilität, die in erster Linie auf eine fehlerhafte Erhöhung der Versuchsgliedervariabilität zurückzuführen ist" (Hühn 1974). (3)

Eine gerade für züchterische Anwendungen (Selektion) wesentliche Folgerung aus diesen Ergebnissen liegt auf der Hand: "Durch die konkurrenzbedingte Vergrößerung der Versuchsgliedervariabilität ergeben sich auch verzernte größere Versuchsgliedermittelunterschiede, die durch die gleichzeitige konkurrenzbedingte Verkleinerung des Versuchsfehlers sich sehr häufig als statistisch signifikant verschieden erweisen, obwohl sie es 'in Wirklichkeit', d.h. bei einer exakteren und umfassenderen Auswertung unter Berücksichtigung der Wirkungen der Konkurrenzerscheinungen, nicht sind. Nichtberücksichtigung dieser Parzellenränderkonkurrenzeffekte, d.h. die in der Praxis meist übliche 'Gesamtparzellenauswertung', führt daher zu einer wesentlich überhöhten Zahl von statistisch signifikanten Versuchsgliedermittelunterschieden" (Hühn 1974).

Auf weitere Ergebnisse und Folgerungen der in Hühn (1974) durchgeführten Konkurrenzuntersuchungen - wobei sämtliche Auswertungen und Berechnungen

anhand der Parzellenmittelwerte vorgenommen wurden! - soll jedoch an dieser Stelle nicht näher eingegangen werden. Bei den nachfolgenden theoretischen Überlegungen wird ausschließlich die Beziehung (1) herangezogen, d.h. in der vorliegenden Arbeit werden theoretische Folgerungen und züchterische Anwendungen unter alleiniger Verwendung dieser Gesamtvarianzen-Ungleichung (1) diskutiert.

Eine Einbeziehung auch noch der in (2) und (3) formulierten Ergebnisse als Voraussetzungen entsprechender Untersuchungen soll Gegenstand einer weiteren Veröffentlichung sein.

2. Einige theoretische Folgerungen unter ausschließlicher Verwendung der Ungleichung (1)

In der üblichen varianzanalytischen Schreibweise läßt sich bei der 'Gesamtparzellenauswertung' für jeden Versuchsort folgendes Untersuchungsmodell formulieren:

$$y_{ijk} = m + a_i + b_j + c_{jk} + d_{ijk} + e_{ijk} \quad (4)$$

Dabei ist:

y_{ijk} = Parzellenmittelwert des i-ten Versuchsgliedes im k-ten Block der Wiederholung j

m = Gesamtmittelwert

a_i = Effekt des i-ten Versuchsgliedes

b_j = Effekt der j-ten Wiederholung

c_{jk} = Effekt des k-ten Blocks in der Wiederholung j

d_{ijk} = Gesamter resultierender Parzellenrand-Konkurrenzeffekt des i-ten Versuchsgliedes im Block k der j-ten Wiederholung

e_{ijk} = Versuchsfehler

Für die Gesamtvarianz V bei der 'Gesamtparzellenauswertung' gilt dann:

$$V = V(y_{ijk}) = V(a_i) + V(b_j) + V(c_{jk}) + V(d_{ijk}) + V(e_{ijk}) + 2 \text{Kov}(a_i, d_{ijk}) + 2 \text{Kov}(b_j, c_{jk}) + 2 \text{Kov}(b_j, d_{ijk}) + 2 \text{Kov}(c_{jk}, d_{ijk}) \quad (5)$$

Hierbei wurde angenommen, daß sowohl der Fehler e_{ijk} mit allen Effekten a_i, b_j, c_{jk} und d_{ijk} als auch a_i mit b_j und auch a_i mit c_{jk} unkorreliert ist.

Bei der 'Parzellenkernauswertung' läßt sich für jeden Versuchsort das folgende Untersuchungsmodell formulieren:

$$y_{ijk}^* = m^* + a_i^* + b_j^* + c_{jk}^* + e_{ijk}^* \quad (6)$$

Hierbei haben $y_{ijk}^*, m^*, a_i^*, b_j^*, c_{jk}^*$ und e_{ijk}^* eine analoge Bedeutung wie die entsprechenden Parameter im Modell (4).

Für die Gesamtvarianz V^* bei der 'Parzellenkernauswertung' erhält man:

$$V^* = V(y_{ijk}^*) = V(a_i^*) + V(b_j^*) + V(c_{jk}^*) + V(e_{ijk}^*) + 2 \text{Kov}(b_j^*, c_{jk}^*) \quad (7)$$

Auch bei dieser Ableitung der Varianzzerlegung (7) wurden wieder analoge Unkorreliertheitsvoraussetzungen wie bei der Berechnung von V in (5) gemacht.

Die beiden Kovarianzterme in der Gleichung (5) - nämlich $\text{Kov}(b_j, d_{ijk})$ und $\text{Kov}(c_{jk}, d_{ijk})$ - können theoretisch zwar existieren, sie dürften jedoch aus verschiedenen Gründen in der Praxis nicht allzu groß sein, weil der Konkurrenzterm d_{ijk} in allererster Linie von dem Versuchsglied abhängen wird und weniger von Block und Wiederholung.

Aufgrund dieser vereinfachenden - aber sicher realistischen - Annahmen erhält man aus (5):

$$V = V(a_i) + V(b_j) + V(c_{jk}) + V(d_{ijk}) + V(e_{ijk}) + 2 \text{Kov}(b_j, c_{jk}) + 2 \text{Kov}(a_i, d_{ijk}) \quad (8)$$

Aus (1), d.h. der Beziehung $V^* \geq V$, folgt mittels (7) und (8):

$$V(a_i^*) + V(b_j^*) + V(c_{jk}^*) + V(e_{ijk}^*) + 2 \text{Kov}(b_j^*, c_{jk}^*) \geq V(a_i) + V(b_j) + V(c_{jk}) + V(d_{ijk}) + V(e_{ijk}) + 2 \text{Kov}(b_j, c_{jk}) + 2 \text{Kov}(a_i, d_{ijk}) \quad (9)$$

Wenn auch nicht unbedingt die einzelnen a_i und a_i^*, b_j und b_j^*, c_{jk} und c_{jk}^* sowie e_{ijk} und e_{ijk}^* jeweils einander gleichgesetzt werden können, so ist

eine realistische Annahme doch sicherlich:

$$V(a_i) = V(a_i^*); V(b_j) = V(b_j^*); V(c_{jk}) = V(c_{jk}^*); \tag{10}$$

$$V(e_{ijk}) = V(e_{ijk}^*); \text{Kov}(b_j, c_{jk}) = \text{Kov}(b_j^*, c_{jk}^*)$$

Mit Hilfe dieser Beziehungen erhält man aus (9):

$$V(d_{ijk}) + 2 \text{Kov}(a_i, d_{ijk}) \leq 0 \tag{11}$$

Eine triviale Folgerung aus (11) ist:

$$\text{Kov}(a_i, d_{ijk}) < 0 \tag{12}$$

Das bedeutet: Je größer der Versuchsgliedeffekt a_i ist, um so kleiner ist der entsprechende Konkurrenzterm desselben Versuchsgliedes - und umgekehrt. Für das spezielle im vorliegenden Experiment mit den Fichtenkreuzungen untersuchte konkurrenzempfindliche Merkmal 'Pflanzenhöhe' erhält man aus den vorliegenden Versuchsergebnissen also das direkt einsichtige und in vielen experimentellen Konkurrenzuntersuchungen nachgewiesene Ergebnis, daß die Pflanzen um so weniger Konkurrenz einflüssen unterliegen je größer sie sind oder umgekehrt: Je kleiner die Pflanzen sind, um so größer wird die entsprechende Konkurrenzkomponente.

Aus (11) ergibt sich noch eine weitere interessante Folgerung: Je größer die Unterschiede in den Konkurrenztermen d_{ijk} sind, d.h. je größer $V(d_{ijk})$, um so stärker negativ ist die Kovarianz zwischen den Versuchsgliedeffekten a_i und den Konkurrenztermen d_{ijk} .

Über den Regressionskoeffizienten $b_{a_i/d_{ijk}}$ des Versuchsgliedeffektes a_i auf den zugehörigen Konkurrenzterm d_{ijk} - dieser Regressionskoeffizient wird im folgenden stets mit b_1 bezeichnet - läßt sich aus (11) eine einfache Größenbeziehung ableiten:

$$b_1 \leq -0,50 \tag{13a}$$

und für den Regressionskoeffizienten b_{d_{ijk}/a_i} des Konkurrenztermes d_{ijk} auf den zugehörigen Versuchsgliedeffekt a_i - dieser Regressionskoeffizient wird im folgenden stets mit b_2 bezeichnet - erhält man aus (11):

$$b_2 \leq -0,50 \cdot f \text{ mit } f = V(d_{ijk})/V(a_i) \tag{13b}$$

Unter Verwendung der bekannten Beziehung zwischen Regressions- und Korrelationskoeffizient lassen sich die drei Parameter b_1 , b_2 und f auf zwei Parameter r und f reduzieren, wobei mit r im folgenden stets der Korrelationskoeffizient zwischen den Versuchsgliedeffekten a_i und den Konkurrenztermen d_{ijk} bezeichnet wird. Die Beziehungen (13a) und (13b) lassen sich dann zusammenfassen in:

$$r \leq -\frac{1}{2} \sqrt{f} \tag{14}$$

Bei einer zusammenfassenden Auswertung aller Versuchsglieder über alle Versuchsorte als Versuchsserie (siehe hierzu: Hühn 1974) erhält man ebenfalls das Ergebnis, daß die Gesamtvarianz bei der 'Parzellenkernauswertung' in allen Untersuchungsjahren größer ist als die Gesamtvarianz bei der 'Gesamtparzellenauswertung'. Hieraus könnte man unter Einbeziehung einer zusätzlichen Komponente für 'Versuchsorte' in die varianzanalytischen Modelle (4) und (6) ähnliche theoretische Untersuchungen und Ergebnisse ableiten, wie es zuvor für den Fall eines einzelnen Versuchsortes geschehen ist. Doch soll auf eine explizite Wiedergabe dieser Resultate an dieser Stelle verzichtet werden.

Abschließend können nun - unter Verwendung der bisher diskutierten Ergebnisse - einige interessante theoretische Folgerungen über die durch bestimmte Regressionsansätze erklärten Anteile an der phänotypischen Gesamtvarianz abgeleitet werden.

Für das Quadrat des Korrelationskoeffizienten r_{PK} zwischen dem phänotypischen Wert und der betreffenden Konkurrenzkomponente erhält man:

$$r_{PK}^2 = [\text{Kov}(y_{ijk}, d_{ijk})]^2 / V(y_{ijk}) \cdot V(d_{ijk}) \tag{15}$$

Für die in (15) auftretende Kovarianz kann man aber schreiben:

$$\begin{aligned} \text{Kov}(y_{ijk}, d_{ijk}) &= \text{Kov}(m + a_i + b_j + c_{jk} + d_{ijk} + e_{ijk}, d_{ijk}) = \\ &= \text{Kov}(a_i, d_{ijk}) + \text{Kov}(b_j, d_{ijk}) + \\ &+ \text{Kov}(c_{jk}, d_{ijk}) + V(d_{ijk}) + \text{Kov}(e_{ijk}, d_{ijk}) \end{aligned} \tag{16}$$

Unter den schon zuvor benutzten Unkorreliertheitsannahmen sowie der Vernachlässigung der beiden Kovarianzterme $\text{Kov}(b_j, d_{ijk})$ und $\text{Kov}(c_{jk}, d_{ijk})$

- Begründung: siehe Ableitung von Formel (8) - erhält man aus (15) und (16):

$$r_{PK}^2 = \frac{[V(d_{ijk}) + \text{Kov}(a_i, d_{ijk})]^2}{V(y_{ijk}) \cdot V(d_{ijk})} \quad (17)$$

Für den durch die Abhängigkeit des phänotypischen Wertes y_{ijk} von der Konkurrenzkomponente d_{ijk} erklärten Anteil v' der phänotypischen Gesamtvarianz $V(y_{ijk})$ folgt aus (17):

$$v' = \frac{[V(d_{ijk}) + \text{Kov}(a_i, d_{ijk})]^2}{[V(d_{ijk})]^2} \cdot \frac{V(d_{ijk})}{V(y_{ijk})} \\ = (1 + b_1)^2 \cdot \frac{V(d_{ijk})}{V(y_{ijk})} \quad (18)$$

Bezeichnet man den Anteil der Varianz der Versuchsgliedeffekte $V(a_i)$ an der Gesamtvarianz $V(y_{ijk})$ mit H - also $H = V(a_i)/V(y_{ijk})$ - so lautet die Beziehung (18):

$$v' = (1 + b_1)^2 \cdot f \cdot H \quad (19)$$

Zusammen mit (13a) ergibt sich nun aus (19) die Abschätzung:

$$v' \leq f/4 \cdot H \quad \text{für } -1,50 \leq b_1 \leq -0,50 \\ v' \geq f/4 \cdot H \quad \text{für } b_1 \leq -1,50 \quad (20)$$

In diesen und den nachfolgenden theoretischen Überlegungen und Formulierungen läßt sich H - d.h. das Verhältnis der Varianz der Sorteneffekte zur Gesamtvarianz - als Maßeinheit benutzen, d.h.: Mißt man v' in H -Einheiten und bezeichnet man dieses so gemessene v' mit v , so folgt aus (20):

$$v \leq f/4 \quad \text{für } -1,50 \leq b_1 \leq -0,50 \\ v \geq f/4 \quad \text{für } b_1 \leq -1,50 \quad (21)$$

Der durch Abhängigkeit des phänotypischen Wertes von der Konkurrenzkomponente erklärte Anteil der phänotypischen Gesamtvarianz beträgt also

- für $-1,50 \leq b_1 \leq -0,50$ höchstens ein Viertel des Varianzanteils dieser Konkurrenzkomponente und
- für $b_1 \leq -1,50$ mindestens ein Viertel des Varianzanteils dieser Konkurrenzkomponente.

Als nächstes interessiert nun noch der durch die Abhängigkeit des phänotypischen Wertes y_{ijk} vom Versuchsgliedeffekt a_i erklärte Anteil der phänotypischen Gesamtvarianz.

Für das Quadrat des Korrelationskoeffizienten r_{PV} zwischen dem phänotypischen Wert und dem betreffenden Versuchsgliedeffekt erhält man:

$$r_{PV}^2 = \frac{[\text{Kov}(y_{ijk}, a_i)]^2}{V(y_{ijk}) \cdot V(a_i)} \quad (22)$$

Für die in (22) auftretende Kovarianz läßt sich nun wieder schreiben:

$$\text{Kov}(y_{ijk}, a_i) = \text{Kov}(m + a_i + b_j + c_{jk} + d_{ijk} + e_{ijk}, a_i) = \\ V(a_i) + \text{Kov}(a_i, d_{ijk}) + \text{Kov}(a_i, e_{ijk}) \quad (23)$$

Wegen der postulierten Unkorreliertheit zwischen Versuchsgliedeffekt und Versuchsfehler folgt somit aus (22) und (23):

$$r_{PV}^2 = \frac{[V(a_i) + \text{Kov}(a_i, d_{ijk})]^2}{V(y_{ijk}) \cdot V(a_i)} \quad (24)$$

Für den durch die Abhängigkeit des phänotypischen Wertes y_{ijk} von dem Versuchsgliedeffekt a_i erklärten Anteil w' der phänotypischen Gesamtvarianz $V(y_{ijk})$ ergibt sich aus (24):

$$w' = \frac{[V(a_i) + \text{Kov}(a_i, d_{ijk})]^2}{[V(a_i)]^2} \cdot \frac{V(a_i)}{V(y_{ijk})} = (1 + b_2)^2 \cdot H \quad (25)$$

Nach einigen einfachen algebraischen Umformungen erhält man aus (25) zusammen mit (13b):

Für $f \leq 2$

$$w' \leq (1-f/2)^2 \cdot H \quad \text{bzw.} \quad w \leq (1-f/2)^2 \\ \text{für } -(2-f/2) \leq b_2 \leq -f/2 \\ w' \geq (1-f/2)^2 \cdot H \quad \text{bzw.} \quad w \geq (1-f/2)^2 \\ \text{für } b_2 \leq -(2-f/2) \quad (26)$$

Für $f \geq 2$

$$w' \geq (1-f/2)^2 \cdot H \quad \text{bzw.} \quad w \geq (1-f/2)^2 \quad \text{für } b_2 \leq -f/2$$

Hierbei ist mit w wieder das in H -Einheiten gemessene w' bezeichnet.

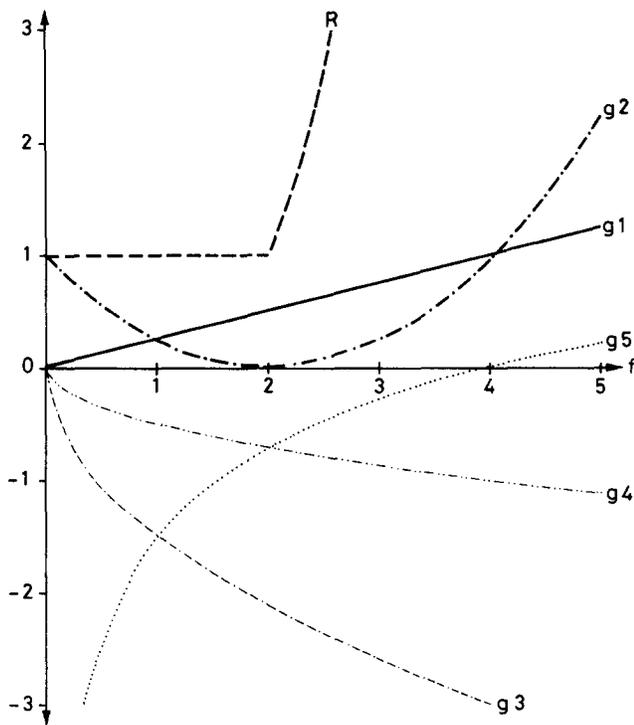


Abb. 1. Graphische Darstellung der Funktionen $g_1(f)$, $g_2(f)$, $g_3(f)$, $g_4(f)$, $g_5(f)$ und $R(f)$ (Erläuterungen: siehe Text)

Der durch Abhängigkeit des phänotypischen Wertes von dem Versuchsgliedeffekt erklärte Anteil der phänotypischen Gesamtvarianz beträgt also

a) für $-(2-f/2) \leq b_2 \leq -f/2$ und $f \leq 2$ höchstens das $(1-f/2)^2$ -fache des Varianzanteils dieser Versuchsgliedeffekte

b) für $b_2 \leq -(2-f/2)$ und $f \leq 2$ mindestens das $(1-f/2)^2$ -fache des Varianzanteils dieser Versuchsgliedeffekte und

c) für $b_2 \leq -f/2$ und $f \geq 2$ mindestens das $(1-f/2)^2$ -fache des Varianzanteils dieser Versuchsgliedeffekte.

Formuliert man die bisher abgeleiteten Ungleichungen (21) und (26) nicht mit Hilfe der Regressionskoeffizienten, sondern mit Hilfe des bereits eingeführten Korrelationskoeffizienten r , so erhält man:

$$v \leq f/4 = g_1(f) \quad \text{für}$$

$$g_3(f) = -1,50\sqrt{f} \leq r \leq -0,50\sqrt{f} = g_4(f)$$

$$v \geq f/4 = g_1(f) \quad \text{für} \quad r \leq -1,50\sqrt{f} = g_3(f)$$

$$w \leq (1-f/2)^2 = g_2(f) \quad \text{für}$$

$$g_5(f) = -(2/\sqrt{f} - \sqrt{f}/2) \leq r \leq -\sqrt{f}/2 = g_4(f)$$

$$\text{und } f \leq 2$$

(27)

$$w \geq (1-f/2)^2 = g_2(f) \quad \text{für}$$

$$r \leq -(2/\sqrt{f} - \sqrt{f}/2) = g_5(f) \quad \text{und } f \leq 2$$

$$w \geq (1-f/2)^2 = g_2(f) \quad \text{für}$$

$$r \leq -\sqrt{f}/2 = g_4(f) \quad \text{und } f \geq 2$$

In der Abb. 1 sind diese Funktionen $g_1(f)$, $g_2(f)$, $g_3(f)$, $g_4(f)$ und $g_5(f)$ graphisch dargestellt.

Aus Abb. 1 lassen sich unschwer die verschiedenen Gültigkeitsbereiche und Größenbeziehungen der zuvor diskutierten Ungleichungen ablesen; dabei erstrecken sich sämtliche Betrachtungen der vorliegenden Arbeit ausschließlich auf das f -Intervall $0 \leq f \leq 4$, wie man z. B. der Ungleichung (14) entnehmen kann.

Für die späteren Überlegungen interessiert besonders die Größenbeziehung zwischen $g_1(f)$ und $g_2(f)$:

$$g_2(f) \leq g_1(f) \quad \text{für } 1 \leq f \leq 4 \quad (28)$$

Wegen $-1 \leq r \leq 0$ gelten die Ungleichungen in (27) zum Teil nur für sehr begrenzte f -Intervalle - und zwar erhält man:

$$\text{Für } 0 \leq f \leq 4 \quad : v \leq g_1(f) \quad \text{für } g_3(f) \leq r \leq g_4(f)$$

$$\text{Für } 0 \leq f \leq 0,44 \quad : v \geq g_1(f) \quad \text{für } r \leq g_3(f) \quad (29)$$

$$\text{Für } 0 \leq f \leq 2 \quad : w \leq g_2(f) \quad \text{für } g_5(f) \leq r \leq g_4(f)$$

$$\text{Für } 2 \leq f \leq 4 \quad : w \geq g_2(f) \quad \text{für } r \leq g_4(f) \quad (30)$$

$$\text{Für } 1,53 \leq f \leq 2,00 \quad : w \geq g_2(f) \quad \text{für } r \leq g_5(f)$$

Es muß an dieser Stelle nochmals besonders darauf hingewiesen werden, daß sämtliche in der vorliegenden Arbeit bis hierher abgeleiteten Ergebnisse nur aus einem einzigen der in den experimentellen Konkurrenzuntersuchungen (Hühn 1974) gefundenen Resultate abgeleitet werden können: Alle bisher in dieser Veröffentlichung erhaltenen Ergebnisse folgen sämtlich nur aus der Beziehung (1): $V^* \geq V$, d. h. aus dem Ergebnis, daß die Gesamtvarianz bei der 'Parzellenkernausswertung' (= Gesamtvarianz ohne Konkurrenz) nicht kleiner ist als die Gesamtvarianz bei der 'Gesamtparzellenauswertung' (= Gesamtvarianz mit Konkurrenz).

Für den in der 'Einleitung' diskutierten Sonderfall $V^* = V$, der - zumindest aufgrund der anfangs zitierten umfangreichen experimentellen Konkurrenzunter-

suchungen - von großer praktischer Bedeutung ist, erhält man aus (11):

$$V(d_{ijk}) + 2 \text{Kov}(a_i, d_{ijk}) = 0 \tag{31}$$

Dieses Ergebnis führt nun zu einer erheblichen Präzisierung der bisherigen Resultate und Abschätzungen:

Für die bereits eingeführten Regressions- und Korrelationskoeffizienten erhält man mittels (31):

$$b_1 = -1/2; \quad b_2 = -f/2; \quad r = -\sqrt{f}/2 \tag{32}$$

und für die durch bestimmte Regressionsansätze erklärten Anteile an der phänotypischen Gesamtvarianz kann man nun schreiben:

$$v' = 1/4 fH \quad \text{bzw.} \quad v = f/4 \tag{33}$$

$$w' = (1-f/2)^2 H \quad \text{bzw.} \quad w = (1-f/2)^2 \tag{34}$$

Aus Abb.1 ($g_1(f) = f/4$ und $g_2(f) = (1-f/2)^2$) lassen sich nun direkt die interessierenden Ergebnisse über die Größenbeziehungen der durch bestimmte Regressionsansätze erklärten Varianzanteile ablesen, z.B.:

$$w \leq v \quad \text{für} \quad 1 \leq f \leq 4 \tag{35}$$

Abschließend soll für diesen Fall $V^* = V$ nun noch der durch einen multiplen linearen Regressionsansatz (phänotypischer Wert y_{ijk} in Abhängigkeit von Versuchsgliedeffekt a_i und Konkurrenzeffekt d_{ijk}) erklärte Teil der phänotypischen Gesamtvarianz untersucht werden.

Für die multiple Bestimmtheit R' erhält man nach bekannten statistischen Formeln (z.B. Sachs 1969):

$$R' = r_{PV}^2 + \frac{(r_{PK} - r_{PV}r)^2}{1 - r^2} \tag{36}$$

Unter Verwendung von $r_{PK}^2 = v' = (1+b_1)^2 fH$ und $r_{PV}^2 = w' = (1+b_2)^2 H$ sowie mit (32) erhält man nach einigen einfachen algebraischen Umformungen:

$$R' = H \quad \text{für alle } f \quad \text{bzw.} \quad R = 1 \quad \text{für alle } f$$

für den Fall, daß die beiden Korrelationskoeffizienten r_{PV} und r_{PK} gleiches Vorzeichen haben, und

$$R' = \frac{2f^2 - 5f + 4}{4 - f} H \quad \text{bzw.} \quad R = \frac{2f^2 - 5f + 4}{4 - f} \tag{37}$$

für den Fall, daß r_{PV} und r_{PK} verschiedenes Vorzeichen haben. (Hierbei wurde wieder das in H-Einheiten gemessene R' mit R bezeichnet.)

Die Aussagen 'gleiche bzw. ungleiche Vorzeichen von r_{PV} und r_{PK} ' lassen sich nun aber in einer sehr einfachen Weise als Bedingungen für f ausdrücken:

$$\begin{aligned} \text{Kov}(y_{ijk}, a_i) &= V(a_i) + \text{Kov}(a_i, d_{ijk}) && \text{(nach(23))} \\ &= V(a_i) - V(d_{ijk})/2 && \text{(nach(31))} \end{aligned} \tag{38}$$

$$\begin{aligned} \text{Kov}(y_{ijk}, d_{ijk}) &= V(d_{ijk}) + \text{Kov}(a_i, d_{ijk}) && \text{(nach(17))} \\ &= V(d_{ijk})/2 && \text{(nach(31))} \end{aligned} \tag{39}$$

Wegen $V(d_{ijk}) \geq 0$ gilt:

"Gleiches Vorzeichen von r_{PV} und r_{PK} " äquiv.
 $"V(a_i) - V(d_{ijk})/2 \geq 0"$ äquiv. " $f \leq 2$ " (40)

"Ungleiches Vorzeichen von r_{PV} und r_{PK} " äquiv.
 $"V(a_i) - V(d_{ijk})/2 \leq 0"$ äquiv. " $f \geq 2$ " (41)

Somit erhält man für die multiple Bestimmtheit R :

$$R = \begin{cases} 1 & \text{für } 0 \leq f \leq 2 \\ \frac{2f^2 - 5f + 4}{4 - f} & \text{für } 2 \leq f < 4 \end{cases} \tag{42}$$

In Abb.1 ist diese Funktion $R = R(f)$ aus (42), die den durch einen multiplen linearen Regressionsansatz erklärten Anteil der phänotypischen Gesamtvarianz wiedergibt, graphisch dargestellt - zusammen mit den Funktionen $g_1(f)$ und $g_2(f)$, die die durch bestimmte einfache lineare Regressionsansätze erklärten Anteile der phänotypischen Gesamtvarianz darstellen.

Für $0 \leq f \leq 2$ ergibt sich somit das interessante Resultat, daß H - d.h. der Anteil der Varianz der Versuchsgliedeffekte an der Gesamtvarianz - gleich ist der multiplen Bestimmtheit R' , die für diesen Bereich also explizit völlig unabhängig von der Größe der Konkurrenzvarianz $V(d_{ijk})$ ist.

In speziellen Fällen (z.B. Klonversuche bei vegetativ vermehrbaren Arten, Feldversuche mit reinen Linien in der Selbstbefruchterzüchtung) und unter bestimmten Voraussetzungen über die Versuchsglied-

effekte a_i stellt H die 'Heritabilität im weiteren Sinn' dar, die für $0 \leq f \leq 2$ also gleich dieser multiplen Bestimmtheit R' ist.

3. Einige Folgerungen für züchterische Anwendungen (positive Massenauslese)

Neben der interessanten Deutung (für $0 \leq f \leq 2$) von H als multipler Bestimmtheit R' lassen sich nun aufgrund derselben Überlegungen einfache Formeln für die Korrelationskoeffizienten r_{PV} und r_{PK} angeben, aus denen sich einige nützliche Folgerungen für züchterische Anwendungen (z.B. positive Massenauslese) ableiten lassen.

Man erhält für r_{PV} und r_{PK} aus (33) und (34):

$$r_{PV} = \left(1 - \frac{f}{2}\right) \sqrt{H} \quad \text{und} \quad r_{PK} = \frac{1}{2} \sqrt{fH} \quad (43)$$

Die Korrelation r_{PK} zwischen dem phänotypischen Wert und der Konkurrenzkomponente ist nach (43) also stets positiv, während die für jede Art von züchterischer Phänotypenauslese wesentliche Korrelation r_{PV} zwischen phänotypischem Wert und Versuchsgliedeffekt sowohl positiv als auch negativ sein kann:

$$r_{PV} \geq 0 \quad \text{für} \quad f \leq 2 \quad \text{und} \quad r_{PV} < 0 \quad \text{für} \quad f > 2 \quad (44)$$

Für $f = 2$, d.h. $V(d_{ijk}) = 2 V(a_i)$, ist $r_{PV} = 0$ - bei beliebigen H -Werten! Dies heißt aber, daß für $f = 2$ ein sinnvolles Schließen vom phänotypischen Wert auf den Versuchsgliedeffekt - und damit züchterisch eine positive Massenauslese - nicht möglich ist. Wegen $f H \leq 1$ kann dieser Fall nur bei H -Werten, die kleiner als 0,50 sind, eintreten. Dieselbe Überlegung mit derselben Schlußfolgerung - nämlich der Unmöglichkeit einer sinnvollen züchterischen positiven Massenauslese - gilt nun aber selbstverständlich nicht nur für $f = 2$ (mit $r_{PV} = 0$), sondern auch für alle f -Werte in einem Intervall um 2, die anhand von N Messungen zu empirisch berechneten r_{PV} -Werten führen, die nicht signifikant von Null verschieden sind bei einer bestimmten vorgegebenen Irrtumswahrscheinlichkeit α . Diese f -Intervalle um 2 werden im folgenden berechnet.

Im Hinblick auf züchterische Anwendungen muß dieser Ansatz jedoch noch erweitert werden: Denn der

bisher diskutierte Fall - r_{PV} nicht signifikant verschieden von $r_{PV} = 0$ - stellt unter züchterischen Gesichtspunkten den allernüchternsten Fall dar, und eine züchterische positive Massenauslese dürfte sich auch noch unter weniger extremen Verhältnissen verbieten. Zu einer sinnvollen quantitativen Formulierung dieser Zusammenhänge kommt man über die Verwendung von Bedingungen über die Bestimmtheitsmaße.

Bezeichnet man mit B_{PV} das Bestimmtheitsmaß des phänotypischen Wertes auf den Versuchsgliedeffekt, so wäre eine sinnvolle Bedingung für die Anwendung der positiven Massenauslese die Forderung eines vorzugebenden zu überschreitenden Mindestwertes $(B_{PV})_{\min}$ für dieses Bestimmtheitsmaß:

$$r_{PV}^2 = B_{PV} > (B_{PV})_{\min} = c \quad (45)$$

Die Frage ist nun: Für welche Intervalle bzw. Tupel (H, N, f, α, c) gilt statistisch signifikant die Ungleichung (45)?

Aus den Konkurrenzbetrachtungen, von denen anfangs ausgegangen wurde, ergeben sich auf diese Art und Weise Folgerungen für die züchterische Selektion - nämlich Zusammenhänge zwischen 'Heritabilität', Größe der Konkurrenzvarianz, notwendiger Mindestanzahl der zu selektierenden Individuen, Irrtumswahrscheinlichkeit und Mindestbestimmtheitsmaß.

Die Beziehungen zwischen diesen Parametern bzw. die zuvor erwähnten züchterisch besonders interessierenden f -Intervalle um $f = 2$ lassen sich nun auf die folgende Weise bestimmen:

Der theoretische Korrelationskoeffizient (in der Grundgesamtheit) zwischen dem phänotypischen Wert und dem Versuchsgliedeffekt ist $r_{PV} = \left(1 - \frac{f}{2}\right) \sqrt{H}$ - und anhand von N Individuen sei dafür der folgende empirische Wert r_{PV} geschätzt worden:

$$r_{PV} = \left(1 - \frac{f}{2}\right) \sqrt{H} + \varepsilon(N) \quad (46)$$

mit $\varepsilon(N) \rightarrow 0$ für $N \rightarrow \infty$.

Nach Anwendung der z -Transformation von R.A. Fisher auf r_{PV} erhält man die folgenden aus der statistischen Literatur bekannten Ergebnisse (siehe z.B. Fisz 1966):

Die Zufallsvariable $z = \frac{1}{2} \ln \frac{1 + r_{PV}}{1 - r_{PV}}$ ist annähernd normalverteilt mit dem Mittelwert

$$\bar{z} = \frac{1}{2} \ln \frac{1 + \rho_{PV}}{1 - \rho_{PV}} + \frac{\rho_{PV}}{2(N-1)}$$

und der Varianz

$$\sigma_z^2 = \frac{1}{N-3}.$$

Dann aber besitzt die Variable $\frac{z - \bar{z}}{\sigma_z}$ eine Standardnormalverteilung

$N(0, 1)$.

Nach Einsetzen und Umformen erhält man somit für die Variable

$$\left(\frac{1}{2} \ln \frac{(1+\rho_{PV})(1-\rho_{PV})}{(1-\rho_{PV})(1+\rho_{PV})} - \frac{\rho_{PV}}{2(N-1)} \right) \cdot \sqrt{N-3} \quad (47)$$

eine $N(0, 1)$ -Verteilung.

Wegen der allgemeinen Forderung $\rho_{PV}^2 > c$, $c \neq 0$ hat man bei positiver Korrelation zwischen phänotypischem Wert und Versuchsgliedeffekt, d.h. $f < 2$, die Nullhypothese $H_0: \rho_{PV} \leq +\sqrt{c}$ gegenüber der Alternativhypothese $H_1: \rho_{PV} > +\sqrt{c}$ zu testen (einseitige Fragestellung), während bei negativer Korrelation zwischen phänotypischem Wert und Versuchsgliedeffekt, d.h. $f > 2$, die Nullhypothese $H_0: \rho_{PV} \geq -\sqrt{c}$ gegenüber der Alternativhypothese $H_1: \rho_{PV} < -\sqrt{c}$ zu testen ist (einseitige Fragestellung).

Für $\rho_{PV} = 0$, d.h. $f = 2$, liegt eine zweiseitige Fragestellung mit der Nullhypothese $H_0: \rho_{PV} = 0$ und der Alternativhypothese $H_1: \rho_{PV} \neq 0$ vor.

Für eine Irrtumswahrscheinlichkeit α werden die Signifikanzgrenzen in der Standardnormalverteilung mit $T(\alpha)$ bezeichnet. Es gilt:

$$\left. \begin{array}{l} T(\alpha) = +1,64 \text{ bzw. } -1,64 \text{ für } \alpha = 5\% \\ T(\alpha) = +2,33 \text{ bzw. } -2,33 \text{ für } \alpha = 1\% \end{array} \right\} \begin{array}{l} \text{bei einseitiger} \\ \text{Fragestellung} \end{array}$$

$$\left. \begin{array}{l} T(\alpha) = \pm 1,96 \quad \text{für } \alpha = 5\% \\ T(\alpha) = \pm 2,58 \quad \text{für } \alpha = 1\% \end{array} \right\} \begin{array}{l} \text{bei zweiseitiger} \\ \text{Fragestellung} \end{array}$$

Bei positiver Korrelation zwischen phänotypischem Wert und Versuchsgliedeffekt, d.h. $f < 2$, erhält man nach Einsetzen von (46) in (47) die folgende Bedingung für die Ablehnung der Nullhypothese, d.h. für die An-

nahme der Alternativhypothese $\rho_{PV} > +\sqrt{c}$ bzw. $\rho_{PV}^2 > c$:

$$\left[\frac{1}{2} \ln \frac{1 + \left(1 - \frac{f}{2}\right) \sqrt{H} + \varepsilon(N)}{1 - \left(1 - \frac{f}{2}\right) \sqrt{H} - \varepsilon(N)} \frac{[1 - \sqrt{c}]}{[1 + \sqrt{c}]} - \frac{\sqrt{c}}{2(N-1)} \right] \times \sqrt{N-3} \geq T(\alpha) \quad (48)$$

Den weiteren Untersuchungen legen wir nun die folgende vereinfachende - aber für die praktische Anwendung sicherlich realistische und ausreichend genaue - Annahme zugrunde: N soll so groß gewählt sein, daß das $\varepsilon(N)$ so klein ist, daß man es mit ausreichender Genauigkeit vernachlässigen kann.

Die Bedingung (48) lautet dann:

$$\left[\frac{1}{2} \ln \frac{1 + \left(1 - \frac{f}{2}\right) \sqrt{H}}{1 - \left(1 - \frac{f}{2}\right) \sqrt{H}} \frac{[1 - \sqrt{c}]}{[1 + \sqrt{c}]} - \frac{\sqrt{c}}{2(N-1)} \right] \cdot \sqrt{N-3} \geq T(\alpha) \quad (49)$$

Einfache algebraische Umformungen der Beziehung (49) führen schließlich zu dem folgenden f -Intervall:

$$2 - \frac{2}{\sqrt{H}} < f \leq 2 - \frac{2}{\sqrt{H}} \cdot \frac{ke^\lambda - 1}{ke^\lambda + 1} \quad (50)$$

Hierin wurde zur Abkürzung gesetzt:

$$\lambda = \frac{\sqrt{c}}{N-1} + \frac{2T(\alpha)}{\sqrt{N-3}} \quad \text{und} \quad k = \frac{1 + \sqrt{c}}{1 - \sqrt{c}}.$$

Im Fall negativer Korrelation zwischen phänotypischem Wert und Versuchsgliedeffekt, d.h. $f > 2$, kommt man aufgrund völlig analoger Überlegungen als Formulierung des Ablehnungsbereiches der Nullhypothese zu dem folgenden f -Intervall:

$$2 + \frac{2}{\sqrt{H}} \cdot \frac{ke^\lambda - 1}{ke^\lambda + 1} \leq f < 2 + \frac{2}{\sqrt{H}} \quad (51)$$

Zusammengefaßt erhält man aus (50) und (51) die Gesamtbedingung für die Annahme der Alternativhypothese - d.h. für die Gültigkeit von $\rho_{PV}^2 > c$:

$$2 + \frac{2}{\sqrt{H}} \cdot \frac{ke^\lambda - 1}{ke^\lambda + 1} \leq f \text{ sowie } f \leq 2 - \frac{2}{\sqrt{H}} \cdot \frac{ke^\lambda - 1}{ke^\lambda + 1} \quad (52)$$

Das für die züchterische Anwendung wesentliche kritische f -Intervall lautet also:

$$U_1 = 2 - \frac{2}{\sqrt{H}} \cdot \frac{ke^\lambda - 1}{ke^\lambda + 1} < f < 2 + \frac{2}{\sqrt{H}} \cdot \frac{ke^\lambda - 1}{ke^\lambda + 1} = U_2 \quad (53)$$

Tab.1. 'Mindestheritabilitäten' H_{\min} für verschiedene Werte von N, c und α

N	c=0,20		c=0,40		c=0,60		c=0,80	
	$\alpha=5\%$	$\alpha=1\%$	$\alpha=5\%$	$\alpha=1\%$	$\alpha=5\%$	$\alpha=1\%$	$\alpha=5\%$	$\alpha=1\%$
25	0,47	0,57	0,65	0,72	0,78	0,83	0,90	0,92
50	0,38	0,46	0,57	0,64	0,73	0,78	0,87	0,89
75	0,35	0,41	0,54	0,59	0,71	0,75	0,86	0,88
100	0,33	0,38	0,52	0,57	0,70	0,73	0,85	0,87
200	0,29	0,33	0,49	0,52	0,67	0,69	0,84	0,85
500	0,25	0,28	0,46	0,48	0,64	0,66	0,83	0,83
1000	0,24	0,25	0,44	0,46	0,63	0,64	0,82	0,82
5000	0,22	0,22	0,42	0,42	0,61	0,62	0,81	0,81

Dies bedeutet nun aber: Alle Wertetupel (H, N, f, c, α), die diese Bedingung (53) erfüllen (unter der Zusatzforderung $fH \leq 1!$), kennzeichnen Situationen, in denen für das entsprechende ρ_{PV} das Bestimmtheitsmaß ρ_{PV}^2 nicht signifikant größer als c ist - also Situationen, in denen eine phänotypische züchterische Selektion (positive Massenauslese) nicht in einer sinnvollen Weise anwendbar sein dürfte.

Für den Fall, daß man für die Anwendbarkeit der züchterischen positiven Massenauslese nur fordert, daß ρ_{PV} signifikant von Null verschieden sein soll, erhält man das entsprechende f-Intervall, indem man in den vorhergehenden Ableitungen und Formeln $c = 0$ setzt (d.h. $k = 1$) und für $T(\alpha)$ die Signifikanzgrenzen der zweiseitigen Fragestellung wählt.

Damit dieses kritische f-Intervall nach (53), in dem eine positive Massenauslese nicht anzuwenden ist, nicht das gesamte f-Intervall $0 \leq f \leq 4$ überdeckt, muß die 'Heritabilität' H gewisse Mindestwerte H_{\min} überschreiten.

Als Bedingung hierfür erhält man:

$$H \geq \left(\frac{ke^{\lambda} - 1}{ke^{\lambda} + 1} \right)^2 = H_{\min} \tag{54}$$

In Tab.1 sind für verschiedene numerische Werte von N, c und α diese 'Mindestheritabilitäten' H_{\min} angegeben. Schon bei relativ kleinen c-Werten liegen diese 'Mindestheritabilitäten' - sogar bei hohen Individuenzahlen N - über 0,50 (siehe Tabelle 1).

Für $H = 0,20$ und $c = 0$ sind in Abb.2 für die Irrtumswahrscheinlichkeiten $\alpha = 5\%$ bzw. $\alpha = 1\%$ die f-Intervalle nach (53), in denen keine positive Massenauslese angewendet werden sollte, dargestellt. Bei kleinen Individuenanzahlen N umfassen diese Intervalle fast den gesamten möglichen f-Bereich $0 \leq f \leq 4$,

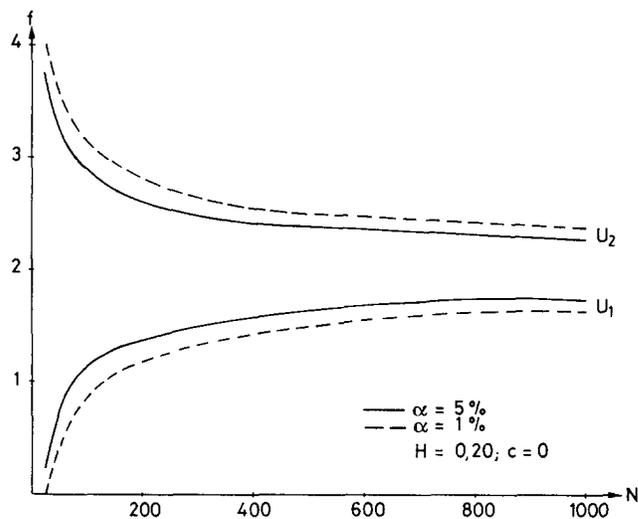


Abb.2. Darstellung der für eine positive Massenauslese nicht zulässigen f-Intervalle für $H = 0,20$, $c = 0$ und unterschiedliche Irrtumswahrscheinlichkeiten α

doch werden für wachsendes N diese Intervalle um $f = 2$ schnell kleiner, betragen jedoch selbst bei $N = 1000$ noch: $1,73 < f < 2,27$ für $\alpha = 5\%$ bzw. $1,64 < f < 2,36$ für $\alpha = 1\%$.

Bei c-Werten ungleich Null sind diese Intervalle selbstverständlich sehr viel ausgedehnter, sie wachsen schnell mit wachsendem c an und nur bei hohen 'Heritabilitäten' H und nicht zu großen c-Werten überdecken sie nicht das gesamte in Frage kommende f-Intervall $0 \leq f \leq 4$, das sich bei größeren H-Werten - völlig unabhängig von den zuvor diskutierten Überlegungen, die zu (53) führten - allein aufgrund der stets zu stellenden Zusatzforderung $fH \leq 1$ weiter stark reduziert. Bei kleinen 'Heritabilitäts-werten' überdecken diese Intervalle nach (53) das gesamte f-Intervall selbst bei hohen Individuenanzahlen N und sogar bei kleinen c-Werten, während bei höheren H-Werten und nicht zu großen Mindestbestimmtheitsmaßen c 'zulässige' f-Werte außerhalb des Intervalls (53) existieren. Doch beschränken sich diese zulässigen f-Werte wegen $fH \leq 1$ ausschließlich auf kleine f-Parameterwerte, d.h. auf Fälle, in denen die Varianz der Versuchsgliedeffekte wesentlich größer als die Varianz der Konkurrenzkomponenten ist: $V(d_{ijk}) \ll V(a_i)$. Das heißt aber, daß eine positive Massenauslese bei größerer Konkurrenzvarianz nicht angewendet werden sollte. Die genauen quantitativen Formulierungen dieser Aussage - in Abhängigkeit von 'Heritabilität', Individuenan-

Tab.2. Für die positive Massenauslese nicht zulässige f-Intervalle (nach (55)) im Sonderfall sehr hoher Individuenzahlen für verschiedene 'Heritabilitätswerte' H und Mindestbestimmtheitsmaße c

H	c=0,20		c=0,40		c=0,60		c=0,80	
	$(U_1)_{N \rightarrow \infty}$	$(U_2)_{N \rightarrow \infty}$						
0,10	0,00	4,00	0,00	4,00	0,00	4,00	0,00	4,00
0,20	0,00	4,00	0,00	4,00	0,00	4,00	0,00	4,00
0,30	0,37	3,33	0,00	3,33	0,00	3,33	0,00	3,33
0,40	0,59	2,50	0,00	2,50	0,00	2,50	0,00	2,50
0,50	0,74	2,00	0,21	2,00	0,00	2,00	0,00	2,00
0,60	0,85	1,67	0,37	1,67	0,00	1,67	0,00	1,67
0,70	0,93	1,43	0,49	1,43	0,15	1,43	0,00	1,43
0,80	1,00	1,25	0,59	1,25	0,27	1,25	0,00	1,25
0,90	1,06	1,11	0,67	1,11	0,37	1,11	0,11	1,11

zahl, Konkurrenzvarianz, Mindestbestimmtheitsmaß und Irrtumswahrscheinlichkeit - liefern die Ungleichungen (53).

Von besonderem Interesse ist noch der Sonderfall sehr hoher Individuenanzahlen. Aus $N \rightarrow \infty$ folgt $\lambda \rightarrow 0$ und das f-Intervall (53) geht über in:

$$(U_1)_{N \rightarrow \infty} = 2 - \frac{2}{\sqrt{H}} \cdot \frac{k-1}{k+1} < f < 2 + \frac{2}{\sqrt{H}} \cdot \frac{k-1}{k+1} = (U_2)_{N \rightarrow \infty} \quad (55)$$

Unter Berücksichtigung der Bedingungen $0 \leq f \leq 4$ und $fH \leq 1$ sind diese f-Intervalle aus (55) in Tab.2 dargestellt. Eine züchterische positive Massenauslese kann in einer sinnvollen Weise also nur für f-Werte mit $f \leq (U_1)_{N \rightarrow \infty}$ angewendet werden. Selbst in diesem Sonderfall sehr hoher Individuenanzahlen beschränkt sich dieser für die positive Massenauslese zulässige f-Bereich also ausschließlich auf kleine f-Parameterwerte. Mit wachsendem c sowie mit abnehmenden 'Heritabilitätswerten' H schrumpft dieser zulässige f-Bereich sehr schnell zusammen (s.Tab.2).

Abschließend soll nun noch die für praktische züchterische Anwendungen wesentliche Frage nach der notwendigen Mindestanzahl N der bei positiver Massenauslese zu selektierenden Individuen diskutiert werden.

Wie die vorhergehenden Untersuchungen gezeigt haben, existiert nur bei hohen 'Heritabilitäten' H sowie nicht allzu großen c-Werten ein für die positive Massenauslese zulässiger sehr schmaler f-Bereich (kleine f-Werte), während in allen anderen Fällen das 'kritische' f-Intervall den gesamten in Frage kommenden Parameterbereich überdeckt - und dieses gilt

in diesen Fällen für beliebige, also sogar auch für besonders hohe, Individuenanzahlen (s.Tab.2).

Aus diesen Gründen ist daher die Frage nach der für die Anwendbarkeit einer positiven Massenauslese mindestens notwendigen Anzahl N der zu selektierenden Individuen in Abhängigkeit von c, H, f und α von keinem allzu großen praktischen Interesse, falls man $c \neq 0$ fordert.

Trotz dieser sich aus den vorhergehenden Überlegungen ergebenden sehr eingeschränkten Anwendbarkeit der positiven Massenauslese, die auch aus vielen aus der Literatur bekannten Untersuchungen hervorgeht, stellt die positive Massenselektion in der Pflanzenzüchtung wegen ihrer einfachen Durchführbarkeit sowie oftmals auch wegen des Fehlens effizienterer Methoden - meist zu Beginn von Züchtungsprogrammen (z.B.: Forstpflanzenzüchtung) - eine sehr häufig angewendete Methode dar, bis sich dann später weitere Züchtungsschritte, z.B. aufgrund von Ergebnissen aus Nachkommenschaftsprüfungen, anschließen.

Folglich wird in der züchterischen Praxis eine positive Massenauslese auch bei einer schwachen Korrelation zwischen phänotypischem Wert und Versuchsgliedeffekt angewendet, d.h. man fordert nur, daß überhaupt eine - wenn auch geringe - signifikante Korrelation besteht.

Nur für diesen Fall, ρ_{PV} signifikant von Null verschieden, sollen nun die notwendigen Mindestanzahlen der zu selektierenden Individuen in Abhängigkeit von H, f und α berechnet werden.

Aus (49) erhält man hierfür die Bedingung:

$$\left| \left(\frac{1}{2} \ln \frac{1 + \left(1 - \frac{f}{2}\right) \sqrt{H}}{1 - \left(1 - \frac{f}{2}\right) \sqrt{H}} \right) \cdot \sqrt{N-3} \right| \geq T(\alpha) \quad (56)$$

Tab.3. Für die Anwendbarkeit der positiven Massenauslese mindestens notwendige Individuenanzahl (berechnet nach (57)) in Abhängigkeit von H, f und α

f	H=0,20		H=0,40		H=0,60		H=0,80	
	$\alpha=5\%$	$\alpha=1\%$	$\alpha=5\%$	$\alpha=1\%$	$\alpha=5\%$	$\alpha=1\%$	$\alpha=5\%$	$\alpha=1\%$
0,00	20	32	10	15	7	10	5	7
0,20	25	40	13	19	9	12	7	9
0,40	31	51	16	25	11	16	8	12
0,60	40	67	20	33	14	22	11	16
0,80	54	91	28	45	19	30	14	22
1,00	78	132	39	66	27	43	20	32
1,20	121	207	61	103	41	68	31	51
1,40	214	369	108	184	72	122		
1,60	481	831	241	415	161	276		
1,80	1922	3327	961	1663				
2,00	-	-	-	-				
2,20	1922	3327	961	1663				
2,40	481	831	241	415				
2,60	214	369						
2,80	121	207						
3,00	78	132						
3,20	54	91						
3,40	40	67						
3,60	31	51						
3,80	25	40						
4,00	20	32						

Diese
Werte ent-
fallen wegen der
Bedingung:
 $fH \leq 1$

wobei für $T(\alpha)$ die Signifikanzgrenzen der zweiseitigen Fragestellung einzusetzen sind. Aus (56) folgt nun durch einfache Umformungen für N:

$$N \geq \frac{4(T(\alpha))^2}{\left[\ln \frac{1 + \left(1 - \frac{f}{2}\right) \sqrt{H}}{1 - \left(1 - \frac{f}{2}\right) \sqrt{H}} \right]^2} + 3 \quad (57)$$

Die für eine sinnvolle Anwendbarkeit der positiven Massenauslese mindestens notwendigen zu selektierenden Individuenanzahlen sind (berechnet nach (57)) in Tab.3 zusammengestellt.

Für ein bestimmtes f nehmen diese Mindestanzahlen mit wachsender 'Heritabilität' H selbstverständlich ab, während sie für einen bestimmten H-Wert bei Annäherung des f-Wertes (von beiden Seiten) an $f = 2$ ansteigen. Für Einzelheiten - z.B. Angaben über die Größe und Schnelligkeit der Mindestanzahlen - Änderungen bei sich ändernden H-Werten bzw. sich ändernden f-Werten - sei auf Tab.3 und Formel (57) verwiesen.

Aus (49) lassen sich nun unschwer auf eine völlig analoge Weise auch die Mindestindividuenanzahlen für c-Werte mit $c \neq 0$ berechnen, die selbstverständlich mit wachsenden c-Werten sehr schnell stark ansteigen. Doch soll an dieser Stelle auf die explizite Wiedergabe der Ergebnisse für $c \neq 0$ verzichtet werden.

Es sei abschließend nochmals besonders darauf hingewiesen, daß die in Tab.3 angegebenen Individuenanzahlen die Mindestanzahlen sind, die man benötigt, damit überhaupt eine signifikant von Null verschiedene Korrelation zwischen phänotypischem Wert und Versuchsgliedeffekt vorliegt. Trotz dieser relativ schwachen Forderung für die Zulässigkeit der Anwendung der positiven Massenauslese dürften die in Tab.3 wiedergegebenen Ergebnisse dem praktischen Züchter nützliche Hinweise und Anwendungsrichtlinien für den anzustrebenden Umfang der Selektion geben.

Literatur

Fisz, M.: Wahrscheinlichkeitsrechnung und Mathematische Statistik. 4. Aufl. Berlin: VEB Deutscher Verlag der Wissenschaften 1966
 Geidel, H.; Haufe, W.: Konkurrenzverhältnisse in Mischbeständen. I. Beschreibung eines Modells mit zwei Komponenten und zwei Nachbarpflanzen als Basis für die Entwicklung differenzierter Modelle. Z. Acker- und Pflanzenbau 127, 269-286 (1968)
 Hühn, M.: Über den Einfluß der Konkurrenz auf die Aussagefähigkeit forstlicher und landwirtschaftlicher Versuche: Experimentelle Ergebnisse aus einer Feldversuchsserie mit Fichtenkreuzungen. Silvae Genetica 23, 77-94 (1974)
 Sachs, L.: Statistische Auswertungsmethoden. Zweite, neubearb. und erw. Auflage. Berlin-Heidelberg-New York: Springer-Verlag 1969

- Singh, K.D.: Vollständige Varianzen und Kovarianzen in Pflanzenbeständen. III. Monte-Carlo-Versuche über den Einfluß der Konkurrenz zwischen Genotypen auf die Voraussage des Ausleseerfolgs. Z. Pflanzenzüchtung 57, 189-253 (1967)
- Stern, K.: Einige Beiträge genetischer Forschung zum Problem der Konkurrenz in Pflanzenbeständen. Allg. Forst- und Jagd-Ztg. 140, 253-262 (1969)

Eingegangen am 23. Oktober 1975/24. Januar 1976
Angenommen durch H. Stubbe

Prof. Dr. M. Hühn
Institut für Pflanzenbau und Pflanzenzüchtung
Universität Kiel
Olshausenstraße 40-60
Neue Universität/Haus S 20a
D-2300 Kiel
Germany (BRD)