

高油品种(含油量为 23.25%)与另一个高油品种(含油量为 23.19%)杂交,育成了含油量为 23.72%(超亲)的哈 70-5071 品系。以另一个以含油量一般的黑农 12 号大豆与一个高油品种哈 64-3519(含油量为 23.97%)杂交,选出了含油量为 23%的哈 70-5179 品系。黑龙江省农科院嫩江农科所用含油量比较高的满仓金大豆(含油量 22%以上)与合丰五号大豆杂交,育成了含油量 22.9%的嫩丰一号大豆。通过以上的结果证实,在杂交选育高油品种时,应以高油品种与高油品种杂交,或者亲本之一必须有一个高油品种。吉林省农科院与山西农科院的研究结果,一致认为,黄大豆、圆形、中大粒(百粒重 20~24.7 克)、淡褐色脐、卵圆形叶、灰白色茸毛、白花的品种含油量高。

(七) 蛋白质含量的亲本选配

大豆蛋白质的遗传方式是受多基因控制

的较复杂的数量遗传性状。杂种后代蛋白含量与亲本关系较大,据研究蛋白质的遗传力为 92%。湖北油料作物所分析了 186 个高蛋白大豆品种,其蛋白质平均含量为 42.98%;其中含量达到 46%以上的有 36 份,有两份超过 50%,最高的一个叫做无名豆,其含量 50.69%。由此可见我国在育成高蛋白品种方面是有条件的。应广泛的搜集高蛋白品种资源,配制高蛋白组合,积极选育高蛋白品种。尽管大豆蛋白质与含油量及籽实产量成负相关,但据日本北海道农业试验场的研究,有的试验材料,油份与蛋白质的合计含量最高值达到 65~67%。这就指出,在不降低或少降低含油量的基础上,可以育成高蛋白品种。为此,在高蛋白亲本的选配中,应选择高蛋白与高蛋白品种杂交,同时还应该选用高蛋白品种与高产品种杂交,使在产量不下降的基础上,育成高产、高蛋白品种。

对变量分析中几个问题的讨论

朱振新

(黑龙江省农科院作物育种所)

在田间试验和统计分析的方法中,随机区组(田间排列)试验、变量分析(统计方法),是被经常采用的试验和分析方法,特别是单因子的品种比较试验或处理效应试验,应用更为普遍。这个试验和分析方法的主要特点是:在随机区组的田间试验中,每个品种或处理在区组(重复)中是随机排列的;每个区组中只设一个对照区;占地面积较少,统计分析时,采用变量分析,可以测算出各种变异原因的变量,包括机误(误差)变量,根据各种变异原因的变量超过机误差变量的倍数大小(即 F 值的大小),来测定各种变量差异的显著性。就是说各种变异原因的变量必须大于机误差变量,并达到一定的倍数时,才

为差异显著。变异原因的变量小于和等于机误差变量,或大于机误差变量但未达一定的倍数,其差异为不显著。这是变量分析的基本原理。

注:变量就是标准差的平方,简化为变量等于差方,习惯称为方差,所以变量分析也叫方差分析。

随机区组、变量分析的试验统计方法,具有试验统计简单,占地面积少,统计分析精确,试验结论可靠等优点。在实际工作中,也发现了一些具体性的问题,根据问题浅谈个人的见解,以供大家讨论。

一、在田间试验时,根据要求和方案,制定了一个完整的随机区组试验,取得了全部试验数据,在统计分析时,选取其中一部

份数据进行分析；或者是在田间淘汰一些品种的处理，把剩下品种的试验数据进行变量分析，是否可以？

田间试验的基本要求是条件一致，也就是要求试验中除了需要比较的那个因素之外，其他一切条件都要完全一致，尽量防止和减少不是试验因子所造成的差异——即试验误差，使数据准确可靠。所以我们希望试验地的条件尽量一致，实际上，如果其他条件不考虑，而试验地的肥力差异总是或大或小地存在的，这种肥力差异就影响试验中的各种数据，干扰了试验处理间的本质上差异，造成了较大的试验误差。因此，我们要求试验地的肥力要均匀，而对试验中的每一个区组范围内的肥力，就更加要求均匀一致。

从统计分析角度来看，如果参加试验的品种（或处理）数多，那么其区组间和区组内的局部控制的范围就大，允许的试验误差也大，而表示变异原因差异显著程度的 F 值标准就小。反之，参加试验的品种数少，其区组间和区组内的局部控制范围就小，允许的试验误差就较小，其 F 值标准就较大。所以，根据上述的推理，可作以下分析：一个完整随机区组的田间试验，如果取这个试验的完整数据作变量分析时，根据机误自由度和区组间或区组内的自由度，查表找出区组间差异显著 $F_{0.05}$ 、 $F_{0.01}$ 值和区组内的差异显著 $F_{0.05}$ 、 $F_{0.01}$ 值。反之，如果只取试验数据中一部分作变量分析时，因为统计分析的品种（或处理）数比原来总数小，这样使其机误自由度和区组内自由度比原来的小，查表所得的区组间和区组内的 $F_{0.05}$ 、 $F_{0.01}$ 值就增大。另外由于机误自由度由大改为了小，在计算品种间均数差异比较显著平准也增大了。由此可见，前者取完整数据作变量分析，是反映出试验的真实性，其分析结果是合理而可靠的。后者是由于种种原因，取其所需的数据作分析，是不符合田间试验的真实情况，提高了变异原因差异显著程度 F 值标准和品种间均数差异比较显著平准，其分析结果是

不够合理而且是不太可靠的。为了阐明问题，举例说明如下：

一个随机区组的品种比较试验，其品种数 13 个，三次重复，其完整数据的变量分析表中的自由度和查表 F 值如下：

变 异 原 因	自 由 度	计 算 F	$F_{0.05}$	$F_{0.01}$
区 组 间	2		3.40	5.51
品 种 间	12		2.18	3.03
机 误	24			
总 数	38			

如果田间淘汰了 4 个品种，把剩下 9 个品种所得的数据进行分析时，其变量分析中的自由度和 F 值如下表：

变 异 原 因	自 由 度	计 算 F	$F_{0.05}$	$F_{0.01}$
区 组 间	2		3.3	6.23
品 种 间	8		2.59	3.09
机 误	16			
总 数	26			

上面两个分析表之间的差异，说明有下列缺点：

1. 第二表虽经数理统计而得的变量分析表，但与田间试验实际情况不符，缺乏真实性。

2. 在变异原因差异显著程度 F 值计算时，第二表中的区组间 $F_{0.05}$ 、 $F_{0.01}$ 和品种间的 $F_{0.05}$ 、 $F_{0.01}$ 值比第一表中相应的 F 值为高，提高了 F 值的标准，这样就容易犯第 II 类错误。这就是说，按实际完整的田间试验取得数据进行分析时，其变异原因差异显著程度的计算 F 值，达显著程度时，而由于取其一部分数据分析时，因 $F_{0.05}$ 、 $F_{0.01}$ 的标准值提高，可能使其计算 F 值不达显著程度。

3. 表中的机误自由度由真实的 24 而变为 16，在进行品种均数差异比较时，如果以 t 测验为例，其品种间均数差异比较显著平准计算如下：

$$\text{相差值 } LSD_{0.05} = \sqrt{\frac{S^2}{N} \times 2} \times t_{0.05}$$

$$\text{相差值 } LSD_{0.01} = \sqrt{\frac{S^2}{N} \times 2} \times t_{0.01}$$

式中 S^2 为机误变量, N 为区组数, 而 $t_{0.05}$ 和 $t_{0.01}$ 均以机误自由度查表所得, 当机误自由度为 24 时, 其 $t_{0.05} = 2.064$, $t_{0.01} = 2.797$; 而当机误自由度为 16 时, 其 $t_{0.05} = 2.120$, $t_{0.01} = 2.921$, 把 t 值代入式中, 使相差值 $LSD_{0.05}$ 、 $LSD_{0.01}$ 也因机误自由度的变化而加大, 标准也提高了, 说明在品种间均数差异比较时也容易犯 II 类错误, 其差异显著结论显然是不太可靠的。

综上所述, 数理统计是农业科学实验的重要研究方法和手段, 为保证试验结论的可靠性而进行统计分析是必要的, 也是必需的。但数理统计一定要符合田间试验的真实情况, 才能使试验结论合理而可靠。

如果在田间试验中, 由于种种原因, 造成缺区时, 可按缺区估计的方法, 估计出缺区的产量, 然后进行变量分析。这样做, 虽然对缺区的估计不十分可靠, 但由于估计了缺区, 进行变量分析时, 其各项自由度、 F 值、机误变量以及计算 $LSD_{0.05}$ 、 $LSD_{0.01}$ 等等, 均是真实的, 取得的结论也是合理而可靠的。

二、增加田间试验中的重复, 是减少土壤差异的主要手段。一般地说重复在试验中的作用有两点:

1. 设置重复能准确地估测试验误差, 即可把试验中总变异分裂为区组间 (重复间) 变异、品种间变异和试验误差三项。不设重复, 就无法估测试验误差。
2. 设置重复可降低试验误差, 提高试验的精确度。因为试验误差 (标准误) 与标准差的关系是:

$$S_{\bar{x}} = \frac{S}{\sqrt{N}}$$

$S_{\bar{x}}$ (标准误) S (标准差) N 为重复次数。
上式说明误差的大小是与重复次数的平方根

成反比。所以重复多, 其试验误差就小。
一般说, 一个试验要求重复 3~5 次即可。一个随机区组的田间试验, 则要求试验中的重复达三次以上为好。实例说明: 一个随机区组品种区域试验, 供试品种为 8 个, 重复三次, 其各项自由度如下表:

变 异 原 因	自 由 度
品 种 间	7
区 组 间	2
机 误	14
总 数	23

如果品种数仍为 8 个, 而重复改为两次, 其各项自由度如下:

变 异 原 因	自 由 度
品 种 间	7
区 组 间	1
机 误	7
总 数	15

由于重复次数减少, 其机误自由度仅为 7, 而机误自由度小于 12 时, 查表所得的 F 值和 t 值标准均偏高, 统计结果取得结论的可靠性就差。如果参加试验的品种为 5 个时, 对其重复三次也嫌少了, 应重复四次为好, 其理由与上相同。所以增加重复, 可使机误自由度增大, 直至大于 12, 使统计结果取得的结论合理而可靠。

三、关于均数间差异比较 (又称多重比较) 的方法: 变量分析中, F 测验只能说明品种间变量是否差异显著。当品种自由度为 1 时, 说明只有两个品种, F 测验结果也具体表明两个品种间的差异是否显著, 不必再另作具体比较。多数的品种试验其品种间自由度是大于 1 的, 所以, 还必须对各品种平均数的差异显著程度进行多重比较。均数差异比较常用的有下列几种方法:

1. t 测验法, 又称 LSD 法或最小显著差数法。是用 $LSD_{0.05}$ 和 $LSD_{0.01}$ 的值来衡量两个品种的均数差异是否显著或极显著。此法

计算简单,应用较为普遍。LSD 法主要适用于两个品种(即品种间自由度等于 1)进行比较。而实际上,许多品种试验的品种数是多于 2 个的(品种间自由度大于 1),如果品种数为 4 个,用一个统一的 $LSD_{0.05}$ 和 $LSD_{0.01}$ 值去比较甲与乙、乙与丙、甲与丙、甲与丁……等之间的均数差异显著程度,显然是不合适的,降低了检验标准的尺度。有时容易犯 I 类错误(所谓 I 类错误是指两个品种间差异实质上是不显著的,由于采用 t 测验法,用统一标准,衡量全体,测验结果使两个品种间差异变为显著)。所以,在品种自由度大于 1 时,采用 t 测验法,在理论上和实践上均是不够合理的。

2. 最小显著极差法(简称 LSR 法)。不同平均数之间的比较采用不同的显著差数标准,克服了 t 测验的局限性。常用的有:① Q 测验法:用不同的 LSR 来衡量不同平均数间的极差。其计算公式:

$$LSR_{0.05} = \sqrt{\frac{S^2}{n}} \times Q_{0.05}$$
$$LSR_{0.01} = \sqrt{\frac{S^2}{n}} \times Q_{0.01}$$

S^2 为机误变量 n 重复数
 $Q_{0.05}$ 和 $Q_{0.01}$ 是以机误自由度和两个平均数间所包含平均数的个数,查 Q 表的值。
②新复极差法又称 SSR 法。此法原理和计算方法与 Q 法相似,只是在计算 $LSR_{0.05}$ 和 $LSR_{0.01}$ 时不是查 Q 表而是查 SSR 表。

其计算公式:

$$LSR_{0.05} = \sqrt{\frac{S^2}{n}} \times SSR_{0.05}$$
$$LSR_{0.01} = \sqrt{\frac{S^2}{n}} \times SSR_{0.01}$$

综上所述:在进行均数间差异比较时,其品种间自由度大于 1 时,采用最小显著极差法(LSR 法)较最小显著差数法(LSD 法)合理而可靠。最小显著极差法中的 Q 法, Q 值偏大,因而提高检验的尺度。新复极差法则纠正了 Q 值偏大,改用 SSR 表。总的说, Q 法和新复极差法的计算都较为复杂而繁琐。目前,我省统一的农作物品种区域试验方案,采用的是 LSDR 法,是既利用了以往习惯的 LSD(即 t 测定法)法,并用 R 值对显著标准进行修正,避免了在品种自由度大于 1 时,采用 LSD 法容易犯 I 类错误的倾向。

亚麻新品种黑亚五号的选育

颜 忠 峰

(黑龙江省农科院经济作物研究所)

亚麻是我省主要经济作物之一,历年生产面积 70~100 万亩,约占全国纤维亚麻面积的 96.0%,亚麻纤维及纺织品除了供应全国和国防军需外,还是重要的传统出口商品,远销世界 90 多个国家和地区。为了适应工农业生产的需要,开创亚麻生产的新局面,我们坚持高产、质优、多抗的育种目标,采用单交、双交及复交的方法,选育出高产稳产、质优、抗旱、抗倒、抗病性强的黑亚五号新

品种,于 1982 年经省农作物品种审定委员会审定推广。

一、选育经过

1963 年以苏联引入的抗立枯病、抗炭疽病强的品种 H-7 为母本,以我国纤维产量高,纤维品质好的华光 1 号为父本,经人工有性杂交后,从杂种后代中选出出麻率高及抗病较好的优系 6302-2。1962 年以苏联的出麻率