table arrangements at Thanksgiving time. S Ct. Dimensions of 2 m by 1.25 m (6 ft by 4 ft).

RABBITEYE BLUEBERRIES (Vaccinium ashei Reade; Vacciniaceae). This vigorous and tall growing shrub makes an excellent hedge when pruned. The waxy, dark green leaves and bright blue to black fruit are very attractive. When kept small, rabbiteye blueberries can be used as a low hedge around the home and as a border or screen plant when unpruned. The blueberry is not without its problems however. Its requirement for acid soils (pH 4.5-5.2), shallow root system, and sensitivity to fertilizers necessitate extra care but the plants are generally long lived. S Ct, Of S. Restricted to dimensions of 2 m by 1 m (6 ft by 3 ft).

TRAILING BLACKBERRY (Rubus trivialis Michx.; Rosaceae). This type of blackberry requires a trellis to allow easier access to fruit and ease of cultivation but are widely adapted and thrive on virtually all soils. Recommended cultivars include 'Oklawaha' and 'Flordagrand' which should be planted together for cross-pollination. S Ct, R Ct, Of S, L. Dimensions 1.5 by 1.5 m (5 by 5 ft).

Proc. Fla. State Hort. Soc. 90:214-215. 1977.

STATISTICAL HYPOTHESIS TESTING: AN ACADEMIC EXERCISE IN FUTILITY

VICTOR CHEW
Agricultural Research Service,
U.S. Department of Agriculture,
217 Rolfs Hall,
University of Florida,
Gainesville, FL 32611

Additional index words. curve fitting, simultaneous interval estimation, selection and ranking procedures.

Abstract. **Statistical hypothesis testing is futile because with enough replications, the null hypothesis will** always **be rejected. More meaningful alternatives to hypothesis testing will be presented.**

If m_1 and m_2 are the unknown true means of 2 treatments, the null hypothesis (denoted by H_0) is that $(m_1 - m_2) = 0$. If H_0 is rejected, it does not mean the difference is big. It only means that the difference is nonzero, which may be very small. If H_0 is accepted, it does not mean that there is no difference, only that the experiment was not large enough. With enough replications, the t-test for comparing 2 treatment means will *always* be highly significant.

Testing the equality of 2 true treatment means is ridiculous. They will always be different, at least beyond the hundredth decimal place. A more sensible hypothesis to test is that $(m_1 - m_2) = d$, a specified minimum difference of practical importance. In a sufficiently large experiment, rejection (or acceptance) of *this* hypothesis implies that the difference *is* large and important (or small and not important). Unfortunately, this generalized t-test cannot be extended to 3 or more treatments. We now discuss several more meaningful alternatives to testing the hypothesis of equal treatment means.

Curve Fitting

If the treatments are levels of a quantitative variable (e.g., temp, time, concn, etc.), the proper statistical technique (e.g., temp, time, conc, etc.), the proper statistical technique is curve fitting or regression analysis (6). We should examine whether the data may be fitted by a straight line (with zero or nonzero slope), a curve that rises and falls, a curve that rises to a maximum and remains constant thereafter, or a curve that approaches but never quite reaches some maximum value (1).

Simultaneous Interval Estimation

The 95% CL (confidence limits) for $(m_1 - m_2)$ are $(\bar{x}_1 - \bar{x}_2) \pm t(.05,n)\sqrt{2s^2/r}$. If the lower limit L is negative and the upper limit U is positive, the null hypothesis H_0 that $(m_1 - m_2) = 0$ will be accepted; if L and U are both negative or both positive, H_0 will be rejected. Interval estimation of $(m_1 - m_2)$ gives an estimate of the difference, beside the information provided by hypothesis testing. Estimation of L and U is the real purpose of the experiment, and not testing the equality of m_1 and m_2. A wide interval (L,U) does not give precise information about $(m_1 - m_2)$. We can calculate the number of replications r such that the confidence interval will be no wider than some preassigned value.

If we use the previous formula to calculate CL for each pair from 3 (or more) means, the probability is individually 95% that $(m_1 - m_2)$, $(m_1 - m_3)$ and $(m_2 - m_3)$ will lie within their respective CL's, but is less than 95% that all 3 differences will *simultaneously* fall between their CL's. A similar multiple comparisons problem arises in testing 3 or more means (2). Tukey's HSD (honestly significant difference) procedure for multiple comparisons of k means is directly applicable to the construction of simultaneous CL's for $k(k-1)/2$ possible pairs of means. The probability is 95% that for *all* pairs (i, j), the difference $(m_i - m_j)$ will lie within $(\bar{x}_i - \bar{x}_j) \pm$ HSD, where HSD = $q(.05; k, n)s/\sqrt{r}$. Values of q(.05; k, n) and q(.01; k, n) are given in (3).

Using the example from Duncan's classic paper, with $k = 7$ means, $r = 6$ replications, $s = 8.92$ with $n = 30$ degrees of freedom, the HSD = $4.46(8.92)/\sqrt{6} = 16.2$. The sample means were $\bar{x}_1 = 71.3$, $\bar{x}_2 = 71.2$, $\bar{x}_3 = 67.6$, $\bar{x}_4 = 61.5$, $\bar{x}_5 = 61.0$, $\bar{x}_6 = 58.1$ and $\bar{x}_7 = 49.6$. The 95% simultaneous CL's for all $7(6)/2 = 21$ pairs of means are of the following form:

$(m_1 - m_7)$: $(71.3 - 49.6) \pm 16.2 = 21.7 \pm 16.2 = (5.5, 37.9)$
$(m_2 - m_6)$: $(71.2 - 58.1) \pm 16.2 = (-3.1, 29.3)$

Note that L will be negative and U will be positive if 2 sample means differ in absolute magnitude by less than the HSD. Other simultaneous estimation procedures are available (7).

Selection and Ranking Procedures

Instead of comparing the k treatments, the objective in some experiments is to (a) pick the best treatment; (b)

select the smallest subset of the k treatments that will have a preassigned probability of at least P of containing the best treatment; (c) pick the t(=2, 3, etc.) best treatments; or (d) rank the t best treatments.

Selecting the Best Treatment

Obviously, the best treatment will be taken to be that one for which the sample mean is the largest, but unless the number r of replications per treatment is sufficiently large, the probability will not be high that the best treatment will give the biggest sample mean. Given a probability P and a difference d of practical importance between the 2 best treatments, we can calculate r such that the probability is at least P that the best treatment will give the largest sample mean. (We have to specify d because if the 2 largest population means are almost equal, it is virtually impossible to ensure a high probability P that the best population will give the largest sample mean; furthermore, no serious loss is incurred in our selecting the second best instead of the best treatment.) We can use the published tables in reverse. Having done the experiment with r replications, we can calculate the probability P that the largest sample mean came from the best population.

Using Duncan's barley example, suppose we wish to have a probability P = .95 that the best population will give the highest sample mean if the 2 best populations differ by at least d = 5 bushels per acre. The formula for r is $r = [h(k, P)\,\sigma/d]^2$. Assume $\sigma = 9$. From tables (4), h(7, .95) = 3.2417, so that the required number of replicates is $r = [9(3.2417)/5]^2 = 35$. The actual experiment was done with r = 6 replications. Solving the above equation, we have $h(7, P) = d\sqrt{r}/\sigma = 5\sqrt{6}/9 = 1.34$. From tables, h(7, P) = 1.34 gives P = .52 approximately.

The above formula assumes that the population standard deviation is known. If this is not available from previous experiments with similar material, sequential and two-stage procedures are available.

Subset selection

If k is large or if the 2 best treatments are close to each other, we may require a prohibitively large experiment to assure a high probability of correctly selecting the best treatment. We may be satisfied instead with picking a small subset for future more intensive study. *We want the smallest* subset such that the probability is P that the best treatment is in the selected subset. The rule is to include a treatment in the subset if its sample mean exceeds $\bar{x}_{max.} - h(k, P)\,\sigma/\sqrt{r}$ (if σ is known) or $\bar{x}_{max.} - \sqrt{2}\,t(1\text{-}p; k\text{-}1, n)\,s/\sqrt{r}$ (σ unknown). Values of t(1-P; k-1, n) are given in Table F1 in (3). In our previous numerical example, $\bar{x}_{max.} = 71.3$, s = 8.92, r = 6, n = 30 and k = 7. If P = .95, the table gives t(.05; 6, 30) = 2.40. The subset includes all those treatments whose means exceed $71.3 - 2.40(8.92)\sqrt{2/6} = 58.94$. The subset containing the 5 largest means has 95% probability of including the best variety. For the other objectives, we refer the reader to (5).

Literature Cited

1. Chew, Victor. 1976. Uses and abuses of Duncan's multiple range test. *Proc. Fla. State Hort. Soc.* 89:251-253.
2. ————. 1976. Comparing treatment means: a compendium. *HortScience* 11:348-357.
3. ————. 1977. Comparisons among treatment means in an analysis of variance. *ARS, USDA, Tech. Bull.* #ARS-H-6 (to appear in March 78).
4. Dudewicz, Edward J. 1976. Introduction to statistics and probability. HRW, New York. (Chapter 11, Ranking and selection procedures.)
5. Gibbons, Jean D., Ingram Olkin and Milton Sobel. 1977. Selecting and ordering populations. Wiley, New York.
6. Petersen, R. G. 1977. Use and misuse of multiple comparison procedures. *Agronomy Journal* 69:205-208.
7. Stoline, M. R. and H. K. Ury. 1977. Tables of the studentized maximum modulus distribution and an application to multiple comparisons among means. (To appear in *Technometrics.*)

Proc. Fla. State Hort. Soc. 90:215-217. 1977.

BREEDING BLUEBERRIES FOR FLORIDA: ACCOMPLISHMENTS AND GOALS[1]

P. M. Lyrene and W. B. Sherman
Fruit Crops Department, IFAS,
University of Florida, Gainesville, FL 32611

Additional index words. Vaccinium, interspecific hybridization, native plants, rabbiteye blueberry.

Abstract. **Ten or more species of wild blueberries are native to Florida. Five species extend as far south as Manatee County, and** Vaccinium myrsinites **Lam. grows south of Miami. Development of a commercial blueberry industry in Florida is favored by the presence of these adapted species from which cultivars can be bred, by the availability of thousands of hectares of acid soils, and by Florida's low latitude, which enables Florida to market blueberries earlier than regions farther north. Accomplishments of the Florida blueberry breeding program to date include identification of low-chilling rabbiteye (V.** ashei **Reade) cultivars that are adapted as far south as Ocala and recent distribution to growers of 3 very early-ripening, low chilling cultivars derived from crosses between northern U.S. highbush cultivars and low-chilling Florida species. Goals of the program include breeding and selection of cultivars for all areas of Florida where blueberry species are native, greater disease resistance, better adaptation to Florida soils and climate, a greater range of ripening dates, and a range of flavors.**

Wild blueberries are widespread in the woods and swamps of Florida. At least 5 of the 10 *Vaccinium* species present in the state (9) are known to occur south of Tampa, and *V. myrsinites* is native in Dade County. The world's first commercial blueberry plantings, consisting of native rabbiteye *(V. ashei)* bushes transplanted from the woods and swamps, were made in north Florida beginning about 1893 (3). By the late 1920's there were an estimated 2,225 acres of cultivated blueberries in Florida (3). Berries were being shipped fresh by rail to northern markets, and at least one

[1]Florida Agricultural Experiment Stations Journal Series No. 806.