United States
Department of
Agriculture

Forest Service

**Southern
Research Station**

Research Paper
SRS-5

Choosing a Transformation in Analyses of Insect Counts from Contagious Distributions with Low Means

W.D. Pepper, S.J. Zarnoch, G.L. DeBarr, P. de Groot, and C.D. Tangren

Choosing a Transformation in Analyses of Insect Counts from Contagious Distributions with Low Means

W.D. Pepper, S.J. Zarnoch, G.L. DeBarr, P. de Groot, and C.D. Tangren

Abstract

Guidelines based on computer simulation are suggested for choosing a transformation of insect counts from negative binomial distributions with low mean counts and high levels of contagion. Typical values and ranges of negative binomial model parameters were determined by fitting the model to data from 19 entomological field studies. Random sampling of negative binomial distributions was simulated and ANOVA's were performed on simulated data for randomized complete block designs with treatment means corresponding to means of negative binomial distributions. The influence of analysis variable, treatment-mean configuration, range in treatment means, significance level of statistical tests, level of contagion, number of blocks, and number of replications in time on observed power and Type I error of F-tests was studied. A computer program was developed to recompute observed power of F-tests for any combination of these factors. The program facilitates choosing a transformation and may also be used to evaluate tradeoffs and determine affordable experimental factors for a future design with a given set of statistical attributes.

Keywords: Contagious distribution, entomology, transformation, variance-mean relationship, variance stabilization.

Introduction

Entomological studies involving the field testing of different treatments require consideration of a number of controllable factors in the planning and analysis phases of the work. In the planning stage, factors common to many studies, such as the number of replicates in space or time and the significance level for tests, must be considered. In some studies, the effect of low insect counts and the skewness of their distribution must also be considered. A common problem with insect counts is that the frequency distribution is frequently skewed to the right, resembling a nonnormal discrete distribution where the variance is related to the mean. Thus, when treatments differ, heterogeneous variance is guaranteed. Parametric analyses of such data tend to produce too many significant results in F-tests and t-tests (Snedecor and Cochran 1967). Therefore, in the analysis phase of a study, two questions must be answered: (1) does the data require transformation, and (2) what type of transformation is required to produce a variate with a more stable variance and a frequency distribution that is approximately normal? The data analyst also decides whether to rank the response from experimental units by treatment and perform nonparametric analyses or perform parametric analyses on the ranks (Conover 1980). Few guidelines are available on how to choose a transformation in parametric analyses of data where insect counts are low or how to determine the influence of various controllable factors on the method of analyses. Thus, our objective was to develop guidelines to help researchers choose a transformation and to develop a simple program[1] that evaluates various tradeoffs and identifies an affordable experimental design with acceptable statistical test attributes.

Methods

Field Data

Preliminary observations on frequency distributions of insect counts were obtained from data sets collected in 19 field studies conducted on various pheromone "treatments" for the white pine cone beetle, *Conophthorus coniperda* (Schwarz), in the United States and Canada (Birgersson and others 1995, DeBarr and others 1995). These pheromone treatments were field tested by installing a trap on a selected tree and baiting it with a preparation (usually chemical formulations); the procedure is replicated in additional trees for each treatment and blocking may or may not be used. Thus, the trap on an individual tree is the experimental or observational unit. The traps are revisited after some period of time and the insects are removed and counted. If the procedure is repeated (usually), the traps are rebaited with the same pheromone treatment and reinstalled on the same tree, or more commonly, new random assignments of treatments are used to reduce tree position bias. The observational response for a treatment is the insect count at the end of the designated time period or the mean count per

[1]The computer program described in this paper is available on request with the understanding that the U.S. Department of Agriculture cannot assure its accuracy, completeness, reliability, or suitability for any other purpose than that reported. The recipient may not assert any propriety rights thereto nor represent it to anyone as other than a Government-produced computer program. The program is written in SAS code but is not a packaged SAS procedure and is not available in other common statistical software packages such as BMDP and Systat. The program (SIMPWR) is available from the senior author upon request: W.D. Pepper, U.S. Department of Agriculture, Forest Service, Southern Research Station, Forestry Sciences Laboratory, 320 Green Street, Athens, GA 30602.

trap where the average is calculated over several time periods. In this paper, the observational response is the mean count per trap per week. The range, average value of insect counts and degree of contagion with respect to counts were observed. Level of contagion is the degree to which the frequencies of high counts and counts near zero exceed those expected in a random process such as the Poisson. The number of pheromone treatments and the number of replications in space and time were observed to determine typical values and ranges (table 1). These field observations were a prerequisite to performing computer simulations comparing alternative methods of analysis on computer-generated data sets with attributes similar to those observed in the 19 field studies.

Statistical Distributions and Transformations

To transform data, its true distribution must be determined. First, a basic statistical model that fits data from a wide range of field conditions should be identified. This model is then used to derive a transformation for analysis. Because ecologists almost exclusively use the negative binomial (NB) model to describe contagious distributions (Kuno 1991), we chose to fit this model to insect counts obtained from the United States-Canadian field studies.

With the NB model, the probability, P_x, that an observational unit will contain x = 0, 1, 2,... insects is

$$P_x = \binom{k+x-1}{k-1} P^x Q^{-(k+x)}, \qquad (1)$$

where

$P = \mu/k$,
$Q = 1+P$,
μ = the population mean, and
k = the contagion parameter.

The NB distribution converges to other distributions as k varies over its range (Anscombe 1949). When k is considerably large, the counts begin to approach randomness and a Poisson model may fit the data. Indeed, as k becomes infinitely large, the NB distribution converges to the Poisson distribution. When k = 1, the NB distribution reduces to the discrete geometric distribution. Values of k in the interval $0 < k < 1$ are associated with high levels of contagion which is characteristic of the NB distribution. As k approaches zero, the count distribution approaches the logarithmic series.

The population mean μ is estimated as the arithmetic average per trap per experimental unit (m) where the average is calculated over replications in time. The variance of individual counts is

$$Var\ (x) = \mu\ (1 + \mu/k)\ . \qquad (2)$$

Thus, variances associated with different treatments with a common value of k are heterogeneous by definition. The moment estimator for Var(x) is the sample variance s^2. The parameter k is estimated with Anscombe's (1949) method 2 which uses successive approximation to choose a value of $\hat{k}$ that makes the following expression approximately equal:

$$f_o = (1 + m/\hat{k})^{-\hat{k}}\ . \qquad (3)$$

In this expression, f_o is the proportion of the total number of traps with a zero count and $(1 + m/\hat{k})^{-\hat{k}}$ is the estimated probability of a zero count computed with the NB model. This is the most efficient estimation method proposed by Anscombe (1949) for estimating a single population value of k when $k < 1$ and $1 < \mu < 10$.

Transformations of counts were discussed in reports published over 60 years ago. Among the first were the square root transformation (Bartlett 1936) and the logarithmic transformation (Williams 1937). These early derivations were based on an assumed linear relationship between the mean (μ) and variance (σ^2) of x, $\sigma^2 = a\mu$, where a is constant. From the linear relationship, a new variate, y = f(x), was derived where the variance of y is constant. Beall (1942) extended this relationship to the nonlinear case,

$$\sigma^2 = \mu + b\mu^2, \qquad (4)$$

which is required to describe some types of field data. This relationship is characteristic of the NB distribution where 1/b = k, the contagion parameter in the NB model. From this nonlinear relationship, Beall (1942) derived the variate,

$$y = b^{-1/2}\ sinh^{-1}(bx)^{-1/2}, \qquad (5)$$

where the variance of y is constant.

Thus, by estimating b in terms of the sample mean and frequency of zero counts and applying the transformation, one obtains the variate y that is theoretically appropriate for use in parametric statistical analyses. It can be shown that the square root and logarithmic transformations are special cases of Beall's (1942) transformation.

Table 1—Statistics from field experiment in the United States and Canada

Test	Treat-ments	Replication in space	Replication in time	m[a]	s^2/m	$\hat{k}$[b]	$Pr > X^2$
1	6	4	3-5	5.7	16.7	0.57	0.27
2	3	10	1	4.0	34.5	.43	.14
3	6	6	1	2.2	4.8	.30	.22
4	5	6	1	1.2	3.8	.32	.87
5	16	8	1-9	0.5	3.2	.26	.001
6	7	8	7	.2	2.3	.20	.85
7a	5	8	6-8	.2	1.4	.19	.14
7b	5	3	1-6	4.5	19.6	.43	.66
7c	5	3	1-6	.4	1.4	.49	.53
7d	5	3	1-6	1.2	3.7	.67	.13
7e	5	3	1-6	.04	1.1	--[c]	--[c]
7f	5	3	1-6	.2	5.5	.06	.28
7g	5	3	1-6	7.0	25.2	.18	.02
8	7	8	6	.4	3.4	.16	.87
9[d]	6	10	2	5.2	9.4	.33	.10
10[d]	5	7	2	2.1	18.7	.18	.70
11[d]	7	6	2	7.2	17.3	.30	.21
13[d]	4	6	2	3.8	13.8	.36	.62
14[d]	4	10	1	1.4	2.9	.57	.04

[a] m is sample mean.
[b] $\hat{k}$ is estimate of contagion parameter.
[c] Every trap had 0 or 1.
[d] Located in Ontario, Canada.

Kuno (1991) cited several comprehensive textbooks and review articles in a recent review of methodology for sampling and analyzing insect populations. Work has continued on the use of variance-mean relationships to describe insect populations. Two types of relationships, empirical and deductive, have been used to describe the variance-mean relationship. One is exemplified by the empirical power equation,

$$\sigma^2 = a\mu^b , \qquad (6)$$

where

a and b = constants.

From this, Bliss (1941) derives the variance stabilizing transformation,

$$y = X^{1-b/2} , \qquad (7)$$

where

b = estimate from data fitted to the variance-mean model.

A second variance-mean relationship is Iwao and Kuno's (1968) deductive quadratic equation,

$$\sigma^2 = A\mu + B\mu^2 , \qquad (8)$$

where

A and B = parameters.

This equation includes Beall's (1942) formula as a special case. The variance stabilizing transformation is also quite similar to Beall's (1942) transformation. Because we were confident that the NB model would fit our data, Beall's (1942) transformation seemed the most appropriate. We saw no need for the generality of the deductive equation of Iwao and Kuno (1968). For comparison, we chose log-transformed counts, the power transformation, ranked values, and untransformed counts as additional analysis variables. Alternative approaches for optimizing power functions are described elsewhere (Bliss 1941, Box and Cox 1964, Perry 1987).

Computer Simulation

A desirable design and analysis controls Type I error and has an acceptable level of power for F-tests of treatments. Our intent was to evaluate alternative designs and transformations with respect to these criteria. However, an analytical solution would be very complex because it would require deriving the distribution of a pseudo F-statistic when the underlying transformation does not have a normal distribution. Thus, we elected to simulate a range of experimental and data-related conditions and compute observed power and Type I error rates for F-tests of treatments.

Fixed and Variable Factors

We used a randomized complete block (RCB) design to compare alternative methods of analyses because this design was used most often in the 19 United States-Canadian studies . The number of treatments was constant at five because this was a typical value. Number of blocks was two, four, or six. It was assumed that insects were collected and empty traps reinstalled on three separate occasions 1 week apart in each simulated experiment and on six occasions 1 week apart in a second replication of these experiments.

Differential treatment effects were produced in simulated studies by generating data for a treatment or a group of treatments from a specified NB distribution.

Random NB counts were generated by first defining a gamma variable (Boswell and Patil 1970) with shape parameter

$$\alpha = k \qquad (9)$$

(the NB contagion parameter) and scale parameter

$$\beta = \mu/k . \qquad (10)$$

where

μ = the mean of the NB distribution.

Using SAS function RANGAM (SAS Institute Inc. 1988a), we generated values of the gamma variable,

$$x = \beta * RANGAM(SEED, \alpha) . \qquad (11)$$

Subsequently, the value of x was used as the mean of a Poisson distribution (Johnson and Kotz, 1969) and SAS function RANPOI was used to generate a random value of

$$y = RANPOI(SEED, x) . \qquad (12)$$

Boswell and Patil (1970) show that y has a NB distribution with mean μ and contagion parameter k . Thus, when selected treatments had different means, so did the NB distribution treatments. The parameter k was varied to include the range of estimates obtained from analyses of the United States-Canadian studies,

$$k = 0.1,\ 0.5,\ and\ 0.75. \qquad (13)$$

The likelihood of a significant F-test depends on the range of treatment means, δ = maximum mean-minimum mean, and the configuration (spacing) of means within the range. In this simulation study, we used treatment-mean configurations described by Young and Young (1991) plus one configuration where all means were equal (fig. 1).

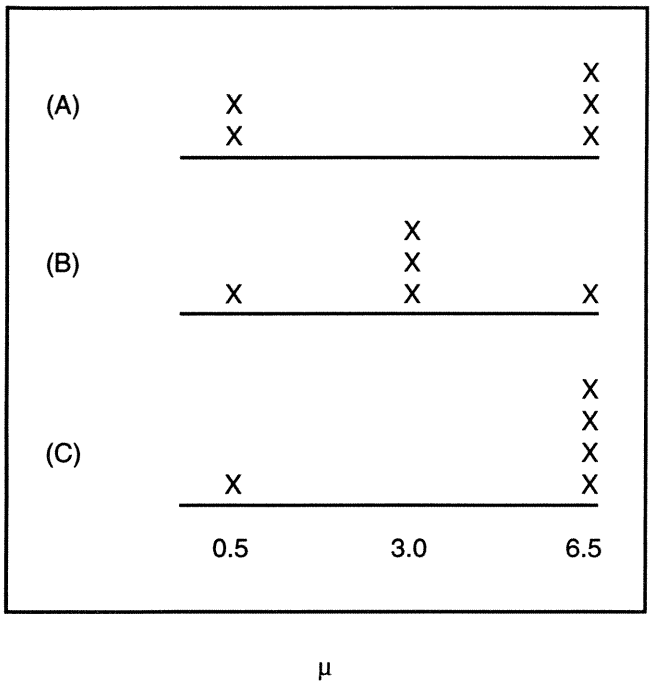


Figure 1—Treatment-mean configurations used in computer simulations. (A) Two subsets of two or more treatments—i.e., treatment means are equal within subsets but different between subsets; number of NB populations = 2. This configuration maximizes the power of the F-test of treatments. (B) All treatment means equal except the smallest and largest which differ from one another; number of NB populations = 3. This configuration minimizes the power of the F-test of treatments. (C) All treatment means equal, except one; number of NB populations = 2. This configuration results in intermediate power.

The smallest mean in each configuration has a value of 0.5. This mean is intended to simulate the value of a control treatment. For a given value of k, the control-treatment mean will have a much smaller variance than a treatment with a mean count of six because the variance of a mean count is

$$Var\,(\bar{x}) = \mu\,(1 + \mu/k)\,/n\,, \qquad (14)$$

where

n = number of replications.

When no contrasts involving the control treatment are planned, some data analysts exclude the control treatment from analysis and reduce overall heterogeneity of variance. However, we elected to judge the performance of selected transformations in the presence of maximum heterogeneity.

In the simulations, each configuration in figure 1 was associated with each of the three values of k to define hypothetical NB populations. Additional populations were defined by using ranges of 2, 4, and 6 and sampled 1,000 times for each RCB design with both 3 and 6 replications in time.

Comparing Transformations in Analyses

Each simulated study data set was subjected to: (1) A parametric ANOVA with SAS procedure GLM (SAS Institute Inc. 1988b) on mean count per trap by treatment and block with means based on untransformed data averaged over both three and six collection dates; (2) parametric ANOVA's on mean count per trap with means based on the transformations,

$$y = sinh^{-1}(\hat{k}/x)^{0.5}(\hat{k})^{0.5} \qquad (15a)$$

$$y = log\,(x + 1) \qquad (15b)$$

$$y = x^{1-b/2} \qquad (15c)$$

where

$\bar{x}$ = the mean count for a given trap,
$\hat{k}$ = the estimated NB contagion parameter, and
b = the estimated parameter for the power function.

and (3) an ANOVA on ranks by block of the mean counts described in (1). We elected to forego classic non-parametric analyses. Friedman's test would be the nonparametric method of choice, but Conover (1980) states that the parametric F-test on ranks performs as well or better than Friedman's test. Our comparisons of analysis variables were based on the ANOVA F-tests. We conducted F-tests at the 1 percent, 5 percent, and 10 percent levels and tallied conclusions as correct or incorrect. When treatment sets were completely homogenous in our simulated studies, we computed an observed Type I error rate by calculating the proportion of F-tests where significance was declared at the nominal level. These observed proportions derived from the five methods of transforming counts were compared. The best analysis variable will yield a proportion that is the smallest and does not exceed the nominal rate.

In simulated studies with differential treatment effects, we computed observed power of F-tests. This is the proportion of F-tests at 1 percent, 5 percent, and 10 percent that correctly reject the null hypothesis and detect treatment differences. Observed power is used as a second criterion for comparing performance of the five analysis variables.

Table 2—Factors studied in simulation experiment

Factors used	Levels of factors
Analysis variables	Y^a RY^b LY^c BY^d PY^e
Significance level for tests (α)	0.01, 0.05, 0.10
Contagion parameter (k)	0.10, 0.50, 0.75
Replication in space (B)	2, 4, 6
Replication in time (R)	3, 6
Range, maximum mean-minimum mean, (δ)	2, 4, 6
Treatment-mean configuration	A, B, C^f

[a] Untransformed counts.
[b] Ranked values of counts.
[c] Logarithmic transformation, log (x+1).
[d] Beall's (1942) transformation, $\sinh^{-1} (\hat{k}/x)^{0.5} (\hat{k})^{0.5}$.
[e] Power transformation, $x^{1-b/2}$.
[f] Configurations A, B, and C are described in figure 1.

We judged the relative performance of each analysis variable and checked for consistency among designs of different size, different levels of contagion in NB populations, different treatment-mean configurations, and different ranges of treatment-mean values and between different numbers of replication in time (table 2).

Results

Modeling Field Data

Table 1 shows estimates of NB population parameters μ and k for the field data. The estimated mean, (m), is defined as the mean insect count per trap averaged over collection dates, block, and treatments. Experimental estimates range from slightly more than seven to less than one. Treatment means (not shown) within experiments range from 14 to less than 1. The low experimental estimates of k, $\hat{k} \leq 0.67$, indicate a high degree of contagion. The high variance-mean ratios are characteristic of the NB distribution.

Goodness-of-fit tests based on Pearson's (1900) chi-square statistic were performed with data from each field test. This test compares observed frequencies in count categories 0,1,2,... with expected frequencies determined by the proposed model. Categories with expected frequencies < 1 are combined (Snedecor and Cochran 1967). The NB model fit well, $0.10 \leq P \leq 0.87$, in 15 of the 19 field studies. Of the remaining four tests, three had $P \leq 0.02$ and the fourth only had two count categories, 0 and 1, prohibiting the goodness-of-fit test.

Computer Simulation Results

Type I Error—Type I errors occurred when treatments in the homogeneous group were declared significantly different. Type I error rates were observed in simulated analyses with each transformation for all combinations of k-values, replication in time, number of blocks, and α-levels of 0.01, 0.05, and 0.10. Results were good for all analysis variables. For all transformations except ranks, the observed Type I error rate never exceeded the nominal value by more than 1 percent. Occasionally, observed rates in analyses of ranked values exceeded the nominal value by 2 percent— still not a serious discrepancy.

Treatment Mean Configuration A—Results for treatment-mean configuration A are shown in figures 2 to 4 for various combinations of k, replications (R) in time and range (δ) in treatment means. Figure 2 (A-F) shows observed power versus number of blocks when testing at the 0.10 level for the highest level of contagion in the data (k = 0.1). When testing at the 0.05 level (k = 0.1), observed power was generally well below 0.80 regardless of required test precision (size of δ) and amount of replication (not shown in figure 2). In these figures, graphs for ranked values, log-transformed values, and Beall's (1942) transformation generally dominate.

When k = 0.1, testing at the 0.10 level still does not result in a satisfactory analysis in most cases. The number of blocks needed to detect a treatment-mean difference of $\delta = 2$ appears to be at least 10.

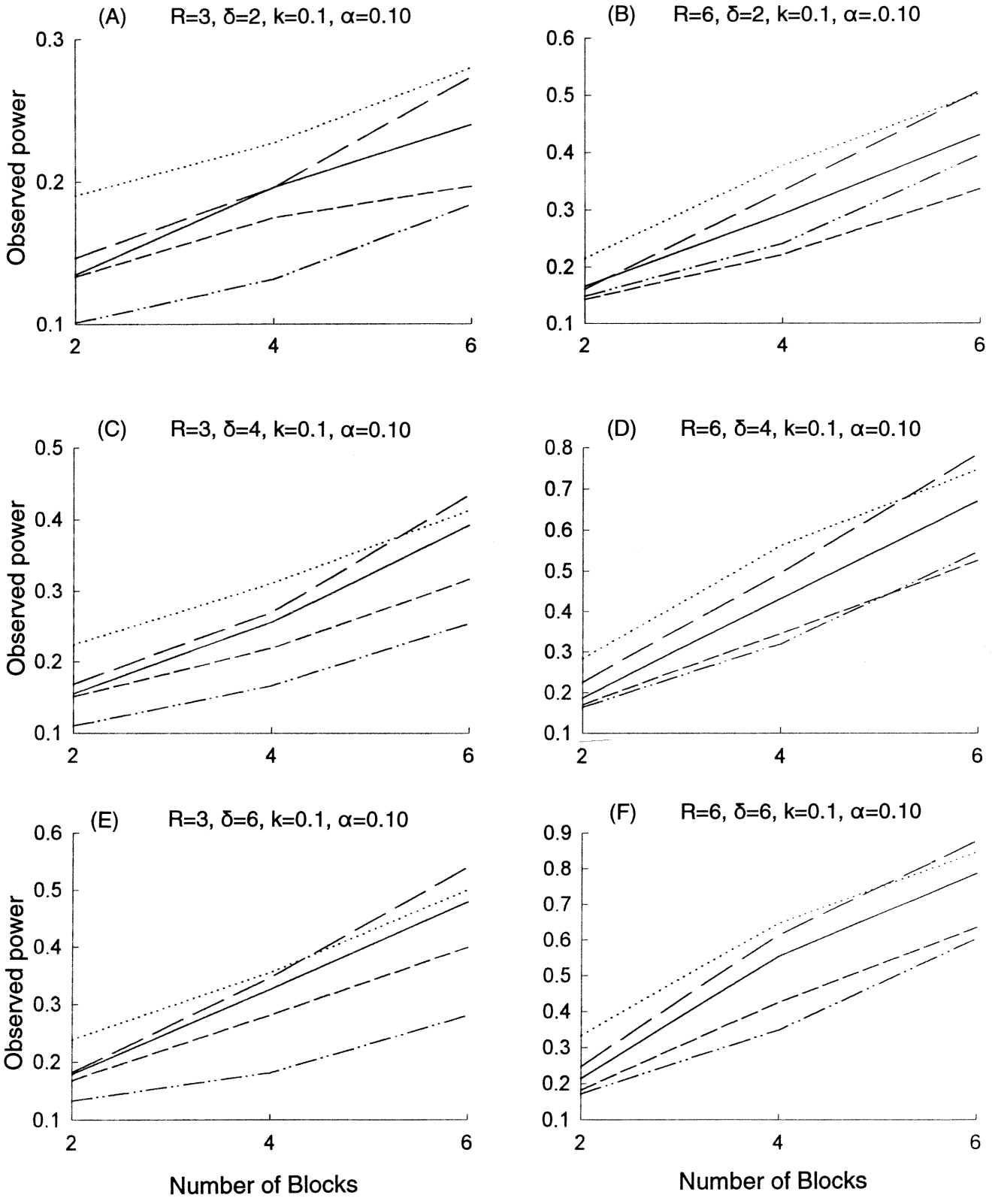


Figure 2 (A-F)—Relationships between observed power of F-tests and number of blocks when treatment-mean configuration = A, k = 0.1, and $\alpha = 0.10$ for five functions of insect counts: ranks (. . . . .), logarithmic transformation (—— ——), power transformation (— — —), Beall's (1942) transformation (——), and the untransformed count (— · · — · ·).

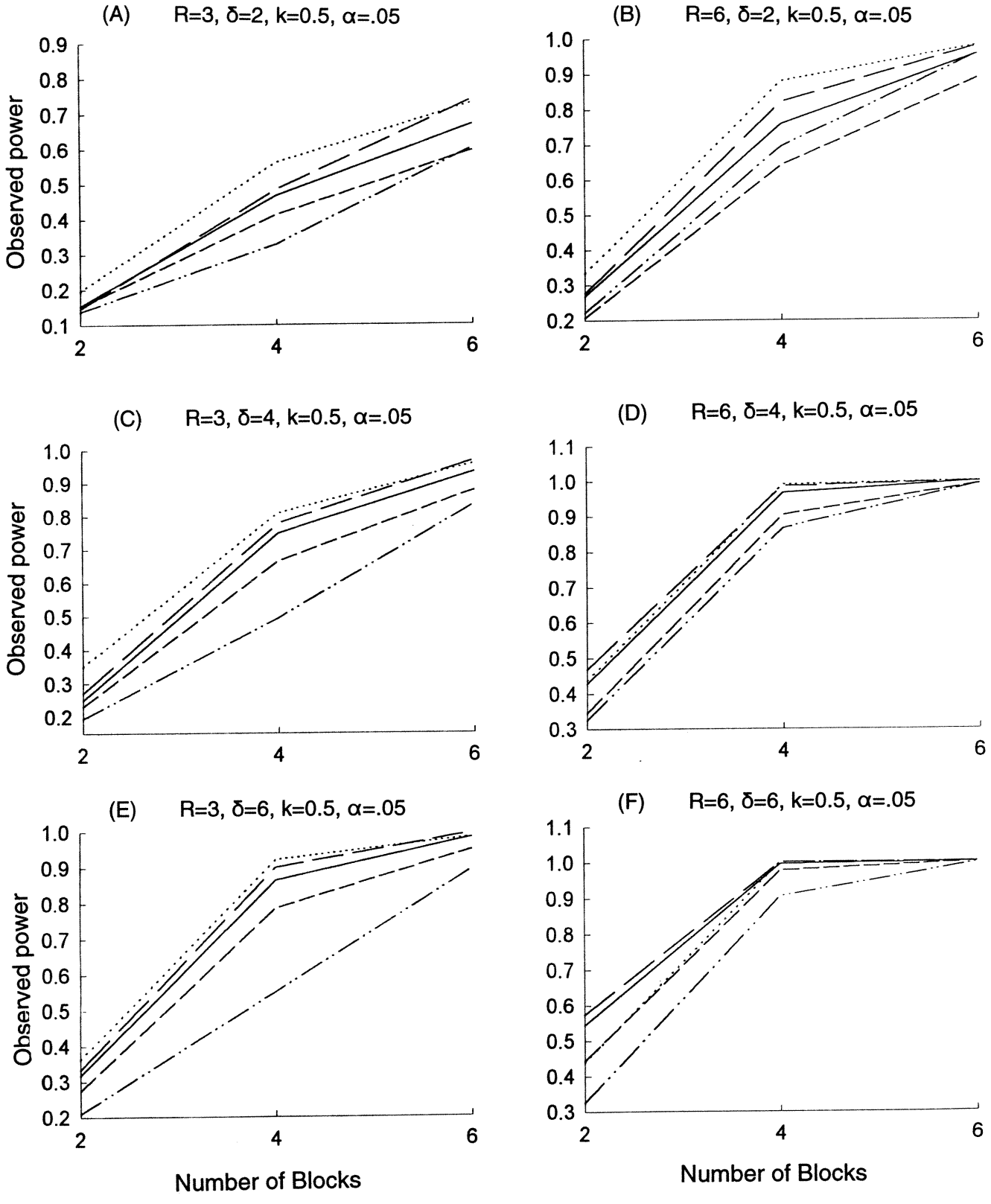


Figure 3 (A-F)—Relationship between observed power of F-tests and number of blocks when treatment-means configuration = A, k = 0.5, and $\alpha = 0.05$ for five functions of insect counts: ranks (. . . . .), logarithmic transformation (—— ——), power transformation (— — —), Beall's (1942) transformation (——), and the untransformed count (— ··—··).

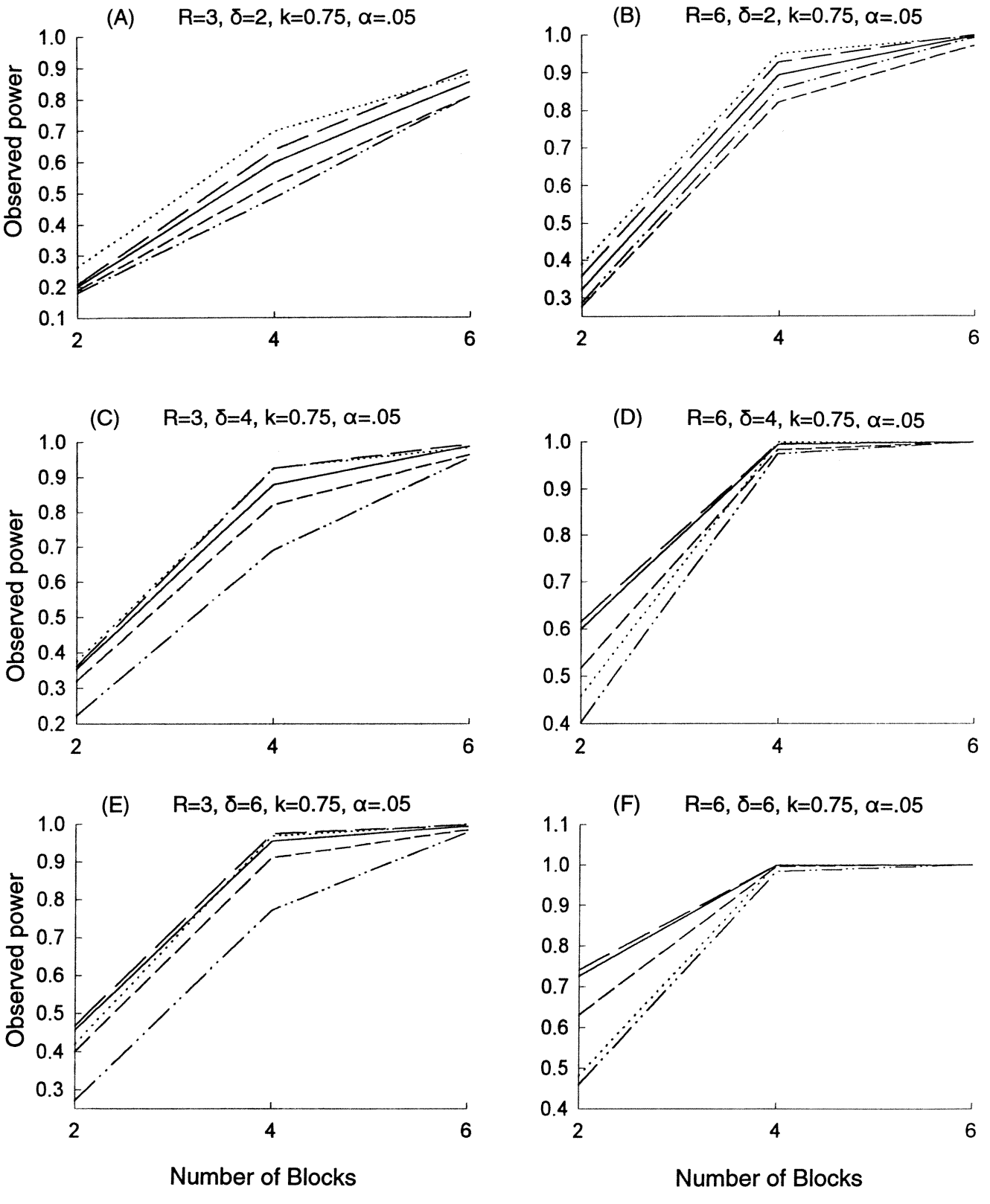


Figure 4 (A-F)—Relationships between observed power of F-tests and number of blocks when treatment-mean configuration = A, k = 0.75, and $\alpha = 0.05$ for five functions of insect counts: ranks (. . . .), logarithmic transformation (—— ——), power transformation (— — —), Beall's (1942) transformation (——), and the untransformed count (— · · — · ·).

When $0.5 \leq k \leq 0.75$ and $\alpha = 0.05$ (figures 3 to 4), four or five blocks are generally needed to keep observed power at or above 0.80 regardless of amount of replication in time and value of δ. Occasionally, three blocks suffice if $\delta \geq 4$ with six replications in time. When testing at the 0.01 level, the same trends are observed, but observed power shifts downward too much.

In most comparisons, analysis variables of choice are ranks and the log transformation. In many cases, the choice between these two variables in terms of power is not clear, and other factors such as relative performance with respect to Type I error and ease of analysis must be considered.

Treatment-Mean Configuration B—Of the three configurations tested, configuration B analyses result in the lowest values of observed power. In figure 1, this is intuitively clear because the range involves only two means resulting in minimal replication. The low power poses a serious problem to the researcher with limited resources.

When $k = 0.1$ and $\alpha = 0.10$, the experiment with the most replication will not provide sufficient power to detect the largest value of δ considered in this study.

When $k = 0.50$ and $\alpha = 0.05$, the largest experiments do not provide sufficient power unless $R > 3$, $\delta > 4$, or both. When $R = 3$, $\delta = 6$, and $B = 6$, observed power is near 0.8 for tests of all analysis variables except the untransformed count. If we increase α to 0.10, observed power can be kept at or above 0.8 with four to six blocks depending on values of R and δ.

When $k = 0.75$ and $\alpha = 0.05$, observed power can be kept above 0.8 with four, five, or six blocks depending on values of R and δ. If $R = 6$ and $\delta = 6$, three blocks will be sufficient.

Treatment-Mean Configuration C—When means are in a C configuration, observed power in analyses of treatments is between maximum-power-A-configuration and minimum-power-B-configuration values. However, when $k = 0.1$ and $\alpha = 0.10$, the experiments with the most replication provided values of observed power below 0.8. When $0.5 \leq k \leq 0.75$ and $\alpha = 0.05$, observed power can be maintained at 0.8 or better with four to six blocks depending on the value of R and δ.

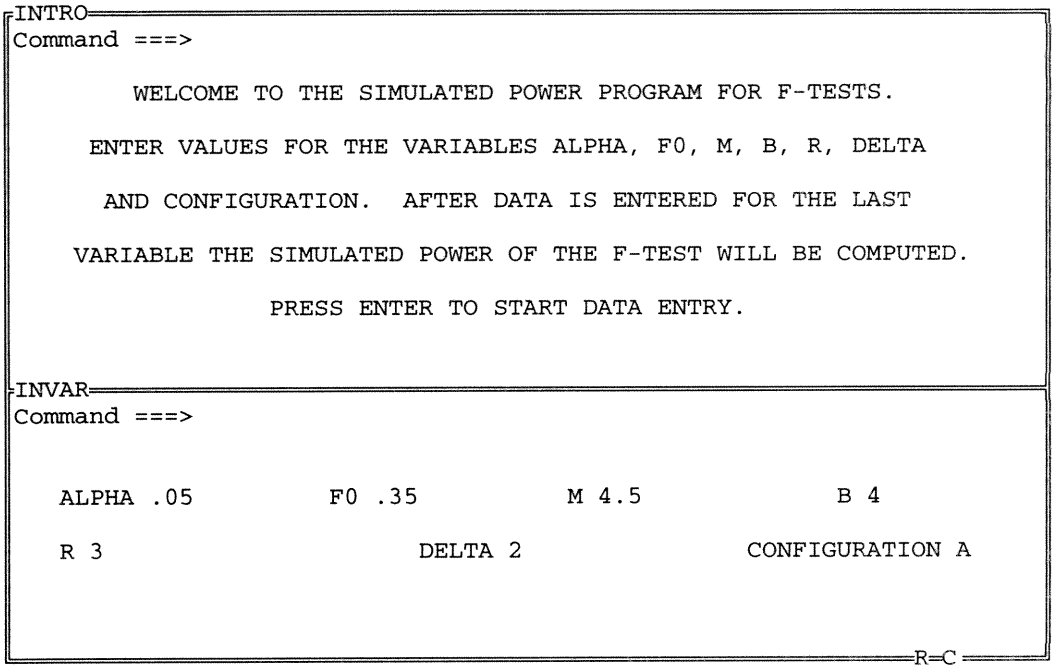
A Numerical Example with the Computer Program

We assume that users will want to know the value of observed power for various combinations of controllable factors. Because our tables and graphs are cumbersome, we developed a partially interactive SAS computer program (SIMPWR) that computes observed power for any combination of experimental factors used in our study. However, this is not a prediction procedure and the user must choose input values only from those used in our simulation studies. The program simply allows the user to determine a tabulated entry of power without handling volumes of computer simulation output.

The following example illustrates the use of SIMPWR with some basic input from Test 7b data, table 1. From table 1, we obtain and input a value of 4.5 for m, the experimental mean. We also input the proportion of traps with zero counts, $f_0 = 0.35$ (not shown). Finally, we must input the values shown in figure 5. In this example, we choose to test at the 5 percent level, $\alpha = 0.05$. The values of m and f_0 are used by the program to estimate $k = 0.43$. The program determines that this estimate is nearer the value of $k = 0.5$ than 0.1 or 0.75. Thus $k = 0.5$ is used in the program. We choose $B = 4$, the number of proposed blocks. Three replications in time are proposed for each treatment, and we wish to detect differences between means as close as two units apart, $\delta = 2$. Furthermore, we assume that true treatment means have the A configuration. Information required by the window shown in figure 5 is entered, and used to calculate the observed power. Both input values and observed power are printed as output. In this case POWER = 0.49 and 0.56 for the logarithmic and rank transformations, respectively. These values are unacceptably low and we must change experimental or test factors to achieve POWER $\geq$ 0.8.

If the design is installed, we are restricted to changes in statistical test attributes. In this second run, we set $\alpha = 0.10$ with additional factors unchanged. This results in POWER = 0.66 and 0.69 for the logarithmic and rank transformations, respectively—still unacceptably low. A third possibility is to set $\delta = 4$ and set other factors at the original values. This results in POWER = 0.78 and 0.81, respectively, for the logarithmic and rank transformations. These values may be regarded as acceptable, but the analysis will be less sensitive in detecting treatment-mean differences.

Suppose we have estimates f_0 and m from a preliminary sample and are at liberty to choose quantity of replication in space and time to control power. We run SIMPWR again with all factors at their originally proposed values except $R = 6$. This results in POWER = 0.82 and 0.88 for the logarithmic and rank transformations, respectively. Linear



SIMULATED POWER PROGRAM FOR F-TESTS
THE INPUT DATA AND FOUR SETS OF OUTPUT

CONFIG	TRANSFRM	ALPHA	FO	M	KHAT	K	B	R	DELTA	POWER
A	RANK	0.05	0.35	4.5	0.43	0.5	4	3	2	0.56237
A	LOG	0.05	0.35	4.5	0.43	0.5	4	3	2	0.48549
A	RANK	0.1	0.35	4.5	0.43	0.5	4	3	2	0.68612
A	LOG	0.1	0.35	4.5	0.43	0.5	4	3	2	0.66366
A	RANK	0.05	0.35	4.5	0.43	0.5	4	3	4	0.80943
A	LOG	0.05	0.35	4.5	0.43	0.5	4	3	4	0.77834
A	RANK	0.05	0.35	4.5	0.43	0.5	4	6	2	0.88025
A	LOG	0.05	0.35	4.5	0.43	0.5	4	6	2	0.82186

Figure 5—Input window and four sets of output from SAS program SIMPWR.

interpolation is possible, but must be used with caution because relationships are not linear. In this example, we may average the values of POWER corresponding to the rank transformation at R =3 and R = 6, (0.81 + 0.88)/2 = 0.845, which corresponds to a value of R between 3 and 6. If POWER = 0.845 is acceptable, R = 5 would be a prudent choice.

Discussion and Conclusions

For a given treatment-mean configuration, observed power of the F-test for a given analysis variable is an increasing function of α, k, number of blocks, number of replications in time, and range in treatment means to be detected. The volumes of graphs and tables produced with computer simulations present a complex picture. This occurs in part because most independent factors interact with at least one other independent factor to exert an influence on observed power. For instance, the effect of increasing number of blocks depends on the level of k (figs. 2 to 4). However, the following generalizations can be made:

- The level of power relationships for all analysis variables depends on treatment-mean configuration. For a given relationship, configuration A produces maximum power. For the same relationship, configuration B produces minimum power and configuration C intermediate power. All three configurations occur in pheromone field tests. However, if a researcher cannot predict which configuration will occur, we recommend using configuration B. This is a conservative approach and should help the researcher avert an overestimation of power for a proposed experiment.

- Regardless of other factors, the analysis variables corresponding to rank, logarithmic, and Beall's (1942) transformations generally dominate the other two with respect to observed power relationships. The transformation of choice is generally between ranked values and log-transformed values and in many cases, choosing will depend on how one views ease of analysis. It is not surprising that the logarithmic transformation has a stabilizing effect on x. The discrete variable x has a frequency distribution which is skewed to the right and resembles the lognormal distribution of the continuous variable z where log (z) has a normal distribution and the estimator for the mean of z is,

$$\hat{\mu}_z = e^{\bar{z} + s^2/2} , \qquad (16)$$

where

$\bar{z}$ = the sample mean, and
s^2 = the sample variance (Johnson and Kotz 1970).

This estimator should be used for the discrete variable x,

$$\hat{\mu}_x = e^{m + s^2/2} , \qquad (17)$$

where

m = a treatment mean, and
s^2 = the ANOVA error mean square on the log scale.

Beall's (1942) transformation was not competitive as a function of k. This transformation will perform best when the true value of k is known, which never occurs in practice. So, even though we knew the value of k for our distributions, we estimated the parameter with our data. We believe this estimation error in k causes Beall's (1942) transformation to be less competitive. In our initial simulations, we inadvertently used the true value of k and Beall's (1942) transformation performed well.

Ranked values competed well for several reasons. These NB distributions have a lot of positive skewness which results in low power for the parametric F-test on raw counts. Skewness is not as much an issue with ranked values because the test statistic is not a function of the raw data, which reflect the magnitude of the skewness directly. Test statistics may be identical with normally distributed data or with highly skewed data as long as the ranks are identical, so power is increased by using the rank transformation, which apparently alleviates the problem of skewness as well as log transformation.

- The level of contagion in the data has a dramatic effect on observed power. This is explained by the influence of k on the variance of individual NB counts for a given mean,

$$Var(x) = \mu(1 + \mu/k) , \qquad (18)$$

and the inverse relationship between power and the variance. For $k = 0.1$, 0.5, and 0.75, the variance of x is respectively $\mu + 10\mu^2$, $\mu + 2\mu^2$, and $\mu + 4\mu^2/3$. When $k \leq 0.1$, maintaining a satisfactory level of power is difficult without making compromises such as increasing the level of significance for testing, increasing δ = maximum mean difference to be detected, or increasing replication.

As shown in the numerical example, SIMPWR can be used to facilitate the choice of analysis for existing data. Furthermore, if preliminary estimates f_0 and m are available, the program can be used to evaluate tradeoffs and determine affordable experimental factors for a future design with a given set of statistical test attributes.

Acknowledgments

We thank John C. Schneider, Department of Entomology and Plant Pathology, Mississippi State University and Floyd Bridgwater, USDA Forest Service, Department of Genetics, North Carolina State University for helpful reviews of an early draft of the manuscript.

Literature Citations

Anscombe, F.J. 1949. The statistical analysis of insect counts based on the negative binomial distribution. Biometrics. 5: 165-173.

Bartlett, M.S. 1936. Square root transformation in an analysis of variance. Journal of the Royal Statistical Society. 3: 68-78.

Beall, G. 1942. The transformation of data from entomological field experiments so that the analysis of variance becomes applicable. Biometrika. 32: 243-262.

Birgersson, G.; DeBarr, G.L.; de Groot, P. [and others]. 1995. Pheromones in white pine cone beetle, *Conophthorus coniperda* (Schwarz) (Coleoptera:Scolytidae). Journal of Chemical Ecology. 21: 143-167.

Bliss, C.I. 1941. Statistical problems in estimating populations of Japanese beetle larvae. Journal of Economic Entomology. 34: 221-232.

Boswell, M.T.; Patil, G.P. 1970. Chance mechanisms generating the negative binomial distributions. In: Patil, G.P., ed. Random counts in scientific work, volume 1. University Park, PA: The Pennsylvania State University Press.

Box, G.E.P.; Cox, D.R. 1964. An analysis of transformations. Journal of the Royal Statistical Society B. 26: 211-252.

Conover, W.J. 1980. Practical nonparametric statistics. 2d ed. New York: John Wiley. 493 p.

DeBarr, G.L.; de Groot, P.; Birgersson, G. [and others]. [In press]. Field bioassays of synthetic pheromones and host monoterpenes for *Conophthorus coniperda* (Coleoptera:Scolytidae). Environmental Entomology.

Iwao, S.; Kuno, F. 1968. Use of the regression of mean crowding on mean density for estimating sample size and the transformation data for the analysis of variance. Researches on Population Ecology. 10: 210-214.

Johnson, N.L.; Kotz, S. 1969. Discrete distributions. Distributions in statistics. New York: John Wiley. 328 p.

Johnson, N.L.; Kotz, S. 1970. Continuous univariate distributions-1. Distributions in statistics. New York: John Wiley. 300 p.

Kuno, Eizi. 1991. Sampling and analysis of insect populations. Annual Review of Entomology. 36: 285-304.

Pearson, K. 1900. On a criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen in random sampling. Philosophical Magazine. 5(50): 157.

Perry, J.N. 1987. Iterative improvement of a power transformation to stabilize variance. Applied Statistics. 36: 15-21.

SAS Institute Inc. 1988a. SAS language guide for personal computers. Release 6.03 ed. Cary, NC: SAS Institute Inc. 558 p.

SAS Institute Inc. 1988b. SAS/STAT user's guide. Release 6.03 ed. Cary, NC: SAS Institute Inc. 1,028 p.

Snedecor, G.W.; Cochran, W.G. 1967. Statistical methods, 6th ed. Ames, IA: The Iowa State University Press. 593 p.

Williams, C.B. 1937. The use of the logarithms in the interpretation of certain entomological problems. Annals of Applied Biology. 24: 404-414.

Young, L.J.; Young, J.H. 1991. Alternative view of statistical hypothesis testing. Environmental Entomology. 20: 1241-1245.

The use of trade or firm names in this publication is for reader information and does not imply endorsement by the U.S. Department of Agriculture of any product or service.

This publication reports research involving pesticides. It does not contain recommendations for their use, nor does it imply that the uses discussed here have been registered. All uses of pesticides must be registered by appropriate State and/or Federal agencies before they can be recommended.

CAUTION: Pesticides can be injurious to humans, domestic animals, desirable plants, and fish or other wildlife—if they are not handled or applied properly. Use all herbicides selectively and carefully. Follow recommended practices for the disposal of surplus pesticides and their containers.

Pepper, W.D.; Zarnoch, S.J.; DeBarr, G.L.; de Groot, P.; and Tangren, C.D. 1997. Choosing a Transformation in Analyses of Insect Counts from Contagious Distributions with Low Means. Res. Pap. SRS-5. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 13 p.

Guidelines based on computer simulation are suggested for choosing a transformation of insect counts from negative binomial distributions with low mean counts and high levels of contagion. Typical values and ranges of negative binomial model parameters were determined by fitting the model to data from 19 entomological field studies. Random sampling of negative binomial distributions was simulated and ANOVA's were performed on simulated data for randomized complete blocks designs with treatment means corresponding to means of negative binomial distributions. The influence of analysis variable, treatment-mean configuration, range in treatment means, significance level of statistical tests, level of contagion, number of blocks, and number of replication in time on observed power and Type I error of F-tests was studied. A computer program was developed to recompute observed power of F-tests for any combination of these factors. The program facilitates choosing a transformation and may also be used to evaluate tradeoffs and determine affordable experimental factors for a future design with a given set of statistical attributes.

Keywords: Contagious distribution, entomology, transformation, variance-mean relationship, variance stabilization.