

1. An oncologist tests the effects of smoking on 4 rats, using an additional 4 rats as controls (non-smoking). The researcher concludes that smoking does not increase the rate of tumor production in lungs, using a criterion of $\alpha = 0.001$. Define Type II statistical error (1), then discuss the oncologist's conclusion with respect to Type II error. (2)

2. Write a complete general linear model corresponding to each of the following tests. Use Y for all response variables. Use X for all explanatory variables on a nominal scale (ie categorical variables). Use Z for all explanatory variables on a ratio scale.

Use subscripts to distinguish and label explanatory variables as needed (e.g. X_1 X_2 etc).

Use the greek letter β (=beta) for parameters that are slopes; use the greek letter γ (=gamma) for parameters that consist of a set of means.

- a. regression (one explanatory variable) (2)

- b. two-way ANOVA, with replication in all cells. (2)

3. If you are interested in hypothesis testing with a statistic whose theoretical frequency distribution is unknown, how can you determine the distribution of this statistic when the null hypothesis is true? (2)

4. Garrod (1967 *Journal of the Fisheries Research Board of Canada* 24:145) computed fishing effort (10^8 vessel ton-hours) and fishing mortality (%/yr) for fish ages 6-7 through 9-10 in the arcto-Norwegian cod fishery. Garrod's data, along with several graphs of the data, are attached on a separate sheet.

a. On the blank line below, write a general linear model to test whether fishing mortality depends upon year. (2)

Be sure to assign symbols to your variables before writing the model. (2)

b. Write the degrees of freedom beneath the appropriate terms in your model. (3)

c. List each symbol in your equation. State whether it is a variable quantity V or a parameter P. State its units. State its dimensions. (10)

| Symbol | V or P ? | Units | Dimensions |
|--------|----------|-------|------------|
|--------|----------|-------|------------|

5. A mycologist takes 3 replicated measurements of the standing stock of soil fungi in each of 3 fields at farm A, then repeats this at farms B and C.

a. Complete the following ANOVA table by filling in the blanks. (4)

| Source | df | Sums of Squares |
|-----------|-------|-----------------|
| farm | _____ | 30.0 |
| field | _____ | 2.3 |
| replicate | _____ | 67.7 |
| total | _____ | |

b. What proportion of the total variability (total Sums of Squares) is due to

farms _____ (1)

fields within farms _____ (1)

replication within fields _____ (1)

c. Which factor (farm, field, or replicate) could be dropped from future studies with the least effect on sensitivity, or ability to detect change in soil fungi over time? (1)

6. Complete the following ANOVA table, for variability in blood pH of 5 strains of mice over 6 days. (10)

| Source | df | SS | MS | F |
|-------------|-------|--------|-------|-------|
| strains | 4 | 0.3680 | _____ | _____ |
| days | 5 | 0.0505 | _____ | _____ |
| interaction | 20 | _____ | _____ | _____ |
| error | _____ | 0.4080 | _____ | |
| Total | 149 | 0.9305 | _____ | |

7. (continued)

c. Provide a complete definition (name, symbol, procedural statement) of an explanatory variable that you can manipulate. (3)

d. Provide a complete definition (name, symbol, procedural statement) of one explanatory variable that you plan to control statistically. (3)

e. Write a general linear model to analyze the results of your experiment. (2)

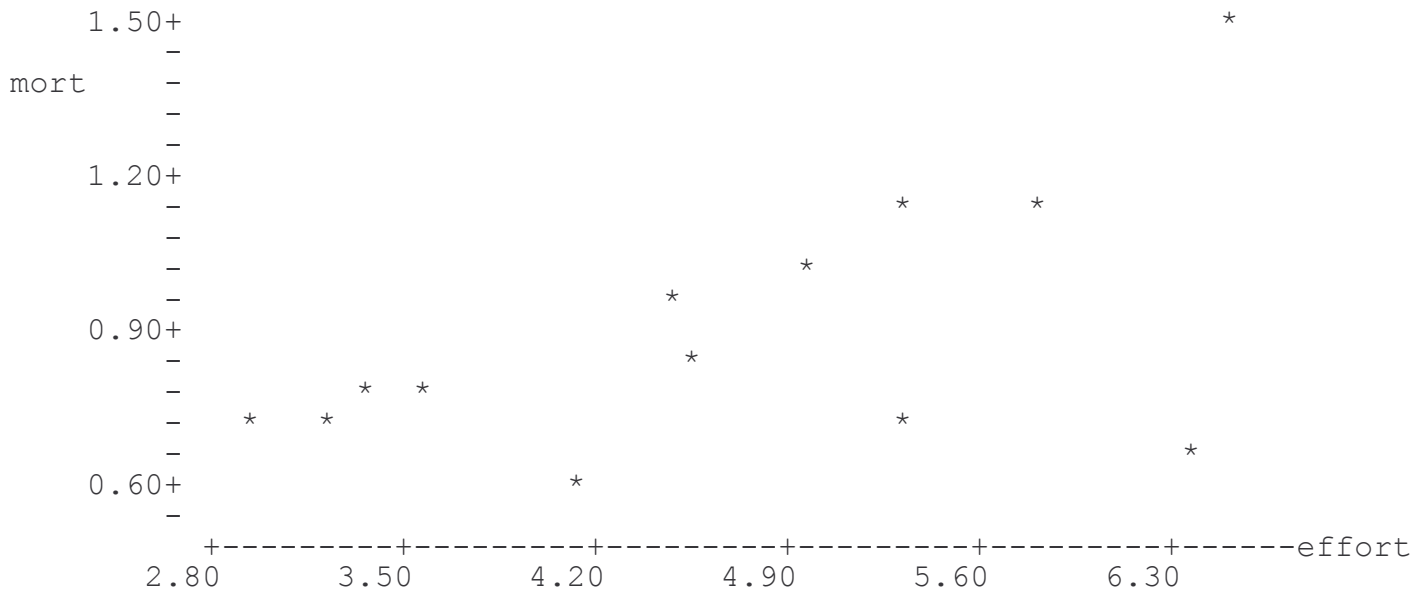
f. State an alternative and null hypothesis pair concerning the treatment variable. (2)

```
MTB > read 'a:garrod.dat' c1 c2 c3;
SUBC> nobs = 13.
      13 ROWS READ
```

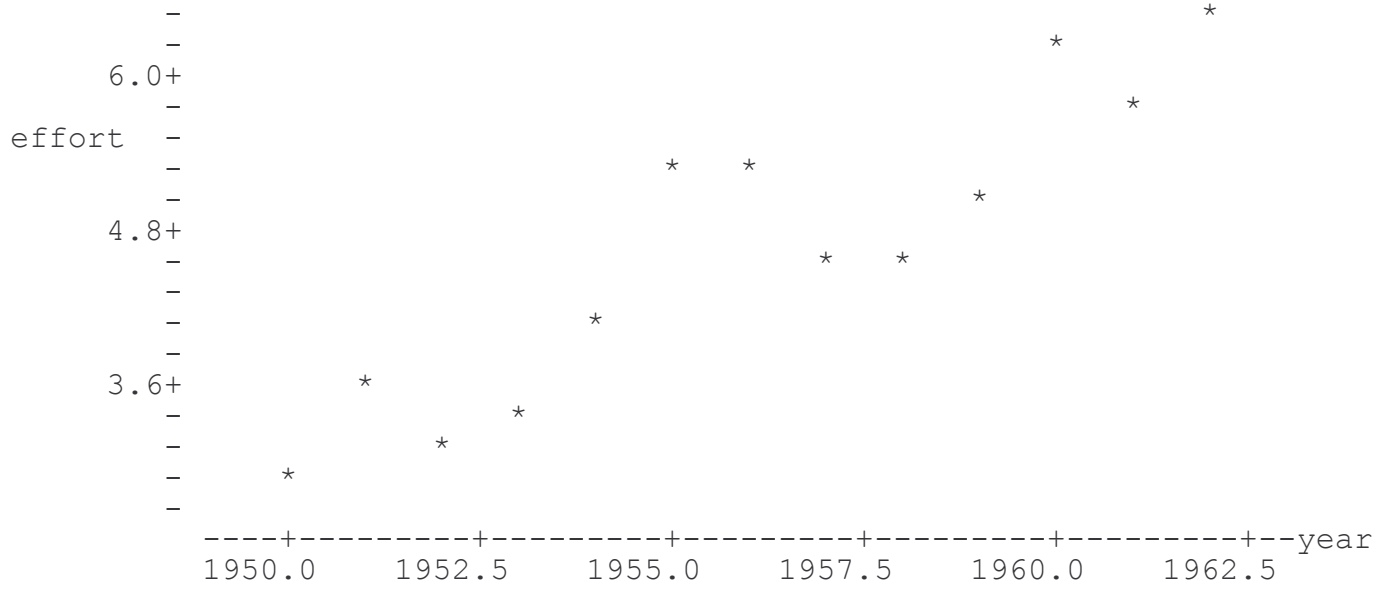
```
MTB > print c1-c3
```

| ROW | effort | mort | year |
|-----|--------|-------|------|
| 1 | 2.959 | 0.734 | 1950 |
| 2 | 3.551 | 0.773 | 1951 |
| 3 | 3.226 | 0.735 | 1952 |
| 4 | 3.327 | 0.759 | 1953 |
| 5 | 4.127 | 0.583 | 1954 |
| 6 | 5.306 | 1.125 | 1955 |
| 7 | 5.347 | 0.745 | 1956 |
| 8 | 4.577 | 0.859 | 1957 |
| 9 | 4.461 | 0.942 | 1958 |
| 10 | 4.939 | 1.028 | 1959 |
| 11 | 6.348 | 0.635 | 1960 |
| 12 | 5.843 | 1.114 | 1961 |
| 13 | 6.489 | 1.492 | 1962 |

```
MTB > name c1 'effort' c2 'mort' c3 'year'
MTB > plot c2 c1
```



MTB > plot c1 c3



MTB > plot c2 c3

