Optional Homework Problems

The attached page 445 is from the Hays (1973) statistics book. The page contains formulas for the density of F, and the mean and the variance of F.

1. Write a SAS or SPSS program to compute the mean and variance of F for any value of numerator, \( \nu_1 \), or denominator, \( \nu_2 \), degrees of freedom.

2. Write a program for the density function of the random variable F. I am not sure whether SAS or SPSS has a factorial function so check for that first. Note that value of the density function is for a given value of F and numerator, \( \nu_1 \) or denominator, \( \nu_2 \) degrees of freedom. The density function will give the height of the F distribution for a given value of F.

3. SAS (Statistical Analysis System) has functions to find an F value given the probability “p”, and numerator “ndf” and denominator degrees of freedom “ddf” and the noncentrality parameter “nc” and a function to find the probability of a certain F with numerator degrees of freedom, denominator degrees of freedom, and the noncentrality parameter. The commands are summarized below:

SAS command to find F for a certain probability.
F=FINV(p,ndf,ddf,nc);

SAS command to find the probability of an F value.
P=PROBF(F,ndf,ddf,nc);

a. Write a program to compute the probability for any value of F.

b. Write a program to compute the F for any p, ndf, ddf, and nc.

c. What does the program below do? 210 is the \( S_{SA} \) and 14.17 is the \( MS_{A} \) for the Keppel one factor ANOVA example that we have used in class.

data retpower;
  noncent=210/14.17;
  fcrit=finv(.95,2,12,0);
  power=1-probf(fcrit,2,12,noncent);
  proc print;
  run;
11.12 TheUse of F Tables

The distribution of $F$ depends upon two parameters $v_1$ and $v_2$.

Since the distribution of $F$ is used to present tables of $F$ distributions, it is difficult to find a table of the results of a problem.

Such tables give only the values of $F$ for the upper proportion of the population.

Note that as the value of $F$ is made increasingly large, the mean of the $F$ distribution approaches 1.00.

Mathematically, the $F$ distribution is derived from the chi-squared distributions, and the distribution can take on other forms when other conditions are set for $v_1$ and $v_2$.

The mean and the variance of an $F$ distribution also depend on the theoretical distribution of $F$ values can be found.

Formally, the density function for the random variable $F$ is defined by

$$f(F; v_1, v_2) = \frac{v_1^{v_1/2} v_2^{v_2/2}}{2^\left(v_1+v_2\right)/2 \Gamma\left(v_1/2\right) \Gamma\left(v_2/2\right)} F^{v_1/2 - 1} (1 + F v_1/v_2)^{-\left(v_1+v_2\right)/2}$$

for $F > 0$.

For $F < 0$ and $v_1, v_2 > 0$ or $F = 0$, the density function is not defined.