Sample size computation activity ST512
("wtloss")

Obtain the sample size necessary to detect differences among weight-reducing agents under certain assumptions.

1. Using SAS code similar to that on p. 115, write a program to compute the sample size necessary to obtain a power of \(1 - \beta = 0.9\) under the assumptions that

\[
\mu_A = \mu_E = 12, \quad \mu_B = 11, \quad \mu_C = 10, \quad \mu_D = 9, \quad \sigma^2 = 1
\]

(a) Create a bunch of values for \(n\), say between \(n = 2\) and \(n = 20\) with a do loop. (Recall the need for an output; statement. Try running without it and you’ll see that the only time writing to the dataset occurs is over the last iteration, when the datastep has reached its end:

```sas
data powercomp;
  do n=2 to 20;
    output;
  end;
run;
proc print data=powercomp;run;
```

(b) Add commands within the do-loop that will create values need to get the appropriate areas underneath \(F\)-distributions:

- numerator and denominator degrees of freedom \(df_1, df_2\)
- the noncentrality parameter \(ncp\)
- critical values \(critval\)

```sas
data powercomp;
  array muvec{5} mu1-mu5 /* create a vector of hypoth. means under H1 */;
  sigma=1;
  do n=2 to 20;
    df1=5-1; df2=(n-1)*5;
    sumtau2=css(of mu1-mu5); /* effect of treatments */
    ncp=sumtau2*n/(sigma*sigma); /* noncentrality ("shift") parameter */
    critval=finv(.95,df1,df2); /* critical value */
    beta = probf(critval,df1,df2,ncp); /* Type II error rate */
    power=1-beta; /* power */
    output;
  end;
  drop mu1-mu5;
run;
proc print data=powercomp;run;
```

(Note the built-in SAS functions: css, finv, probf.)
(c) Plot the power against the sample size.

```sas
proc gplot data=powercomp;
plot power*n;
run;
```

(d) Enhance the plot with a `symbol` statement before the `gplot`:

```sas
symbol value=dot;
proc gplot data=powercomp;
plot power*n;
run;
```

(e) For the homework, you’re asked about how things change as $\sigma$ increases. Make an outside do loop with three values of $\sigma = 1, 1.5, 2.0$:

```sas
data powercomp;
array muvec(5) mu1-mu5 (12,12,11,10,9); /* creates a vector of hypoth. means under H1 */
sigma=1;
do sigma=1 to 2 by 0.5;
do n=2 to 20;
   df1=5-1; df2=(n-1)*5; /* effect of treatments */
   sumtau2=css(of mu1-mu5); /* noncentrality ("shift") parameter */
   ncp=sumtau2*n/(sigma*sigma); /* critical value */
   critval=finv(.95,df1,df2); /* Type II error rate */
   beta = probf(critval,df1,df2,ncp); /* power */
   power=1-beta;
   output;
end;
end;
drop mu1-mu5;
run;
```

(f) Produce a different symbol for each value of $\sigma$:

```sas
symbol value=dot;
proc gplot data=powercomp;
plot power*n;
run;
```

(g) Connect the dots with the `join` option for the `interpolate` command in the `symbol` statement:

```sas
symbol value=dot i=join;
proc gplot data=powercomp;
title "HW3, problem 2";
plot power*n;
run;
```