STAT 350: Lecture 29

Power and Sample Size Calculations

Definition: The power *function* of a test procedure in a model with parameters θ is P_θ(Reject).
Definition: The non-central distribution t with non-centrality parameter δ and degrees of freedom ν is the distribution of

$$\frac{Z}{\sqrt{U/\nu}}$$

where Z is $N(\delta, 1)$, U is χ^2_ν and Z and U are independent When $\delta = 0$ we get the usual, or **central** t distribution.
Fact: If a is a vector of length p and a_0 is some scalar then

$$\frac{a^T\hat{\beta} - a_0}{\sqrt{MSE}\sqrt{a^T(X^TX)^{-1}a}}$$

has a non-central t distribution with non-centrality parameter

$$\delta = \frac{a^T\beta - a_0}{\sigma\sqrt{a^T(X^TX)^{-1}a}}\,.$$

Power of two sided tests from table B 5. Normally computed before experiment based on assumptions about a_0, σ and X^TX.

Sample Size determination

Before an experiment is run it is sensible, if the experiment is costly, to try to work out whether or not it is worth doing. You will only do an experiment if the probability of Type I and II errors are both reasonably low. The simplest case arises when you prespecify a level, say $\alpha = 0.05$ and an acceptable probability of Type II error, say 0.10, for testing a null hypothesis like $a^T\beta = a_0$. Then you need to specify

- The ratio $(a^T\beta - a_0)/\sigma$; this value comes from a physically motivated understanding of what value of the discrepancy $a^t\beta - a_0$ would be important to detect and from some understanding of the roughly what values might be reasonable for σ.

- How the design matrix would depend on the sample size. The easiest thing is to fix some small set of say j values $x_1, \ldots, x_j$ and then use each member of that set say m times so that the aggregate sample size is mj. This gives a non-centrality parameter of the form

$$\frac{a^T\beta - a_0}{\sigma} \times \frac{\sqrt{m}}{\sqrt{a^T(X^TX)^{-1}a}}$$

  The value $n = mj$ influences both the row in table B.5 which should be used and the value of δ. If the solution is large, however, then all the rows in B.5 at the bottom of the table are very similar so that effectively only δ depends on n; we can then solve for n.

F tests

Simplest example: regression through the origin (no intercept term.)

- Model

$$Y_i = \beta_1 X_{i,1} + \cdots + \beta_p X_{i,p} + \epsilon_i$$

- Test $\beta_1 = \cdots = \beta_p = 0$
- F statistic

$$F = \frac{MSR}{MSE} = \frac{\hat{Y}^T\hat{Y}/p}{\hat{\epsilon}^T\hat{\epsilon}} = \frac{Y^THY/p}{Y^T(I-H)Y/(n-p)}\,.$$

Suppose now that the null hypothesis is false.

- Substitute $Y = X\beta + \epsilon$ in F.
- Use $HX = X$ (and so $(I-H)X = 0$).
- Denominator is

$$\frac{\epsilon^T(I-H)\epsilon}{n-p}$$

- So: even when the null hypothesis is false the denominator divided by σ^2 has the distribution of a χ^2 on $n-p$ degrees of freedom divided by its degrees of freedom.

- FACT: Numerator and denominator are independent of each other even when the null hypothesis is false.
- Numerator is
$$\frac{(\epsilon + X\beta)^T H(\epsilon + X\beta)}{p}$$
- Divide by σ^2 and rewrite this as
$$W^T H W/p$$
- $W = (\epsilon + X\beta)/\sigma$ has a multivariate normal distribution with mean $X\beta/\sigma = \mu/\sigma$ and variance the identity matrix.

FACT:

If W is a $MVN(\tau, I)$ random vector and Q is idempotent with rank p then $W^T QW$ has a **non-central** χ^2 distribution with non-centrality parameter

$$\delta^2 = E(W^T QW) - p = \tau^T Q\tau$$

and p degrees of freedom. This is the same distribution as that of

$$(Z_1 + \delta)^2 + Z_2^2 + \cdots + Z_p^2$$

where the Z_i are iid standard normals. An ordinary χ^2 variable is called **central** and has $\delta = 0$.

FACT:

If U and V are independent χ^2 variables with degrees of freedom ν_1 and ν_2, V is central and U is non-central with non-centrality parameter δ^2 then

$$\frac{U/\nu_1}{V/\nu_2}$$

is said to have a **non-central** F distribution with non-centrality parameter δ^2 and degrees of freedom ν_1 and ν_2.

POWER CALCULATIONS

- Table B 11 gives powers of F tests for various small numerator degrees of freedom and a range of denominator degrees of freedom
- Must use $\alpha = 0.05$ or $\alpha = 0.01$.

- In table ϕ is our $\delta/\sqrt{p+1}$ (that is, the square root of what I called the non-centrality parameter divided by the square root of 1 more than the numerator degrees of freedom.)

SAMPLE SIZE CALCULATIONS

- Sometimes done with charts and sometimes with tables; see table B 12.
- This table depends on a quantity

$$\frac{\Delta}{\sigma} = \sqrt{\frac{(p+1)\delta^2}{n}}$$

  To use the table you specify

  - α (one of 0.2, 0.1, 0.05 or 0.01)
  - Power ($= 1 - \beta$ in notation of table)– must be one of 0.7, 0.8, 0.9 or 0.95
  - Non-centrality per data point, δ^2/n.

  Then you look up n.

- Realistic specification of δ^2/n **difficult** in practice.

Examples

POWER of t test: SAND and FIBRE example.

Consider fitting the model

$$Y_i = \beta_0 + \beta_1 S_i + \beta_2 F_i + \beta_3 F_i^2 + \epsilon_i$$

Compute power of t test of $\beta_3 = 0$ for the alternative $\beta_3 = -0.004$. (This is roughly the fitted value. In practice, however, this value needs to be specified *before* collecting data so you just have to guess or use experience with previous related data sets or work out a value which would make a difference big enough to matter compared to the straight line.)

Need to assume a value for σ. I take 2.5 – a nice round number near the fitted value. Again, in practice, you will have to make this number up in some reasonable way.

Finally $a^t = (0, 0, 0, 1)$ and $a^T(X^TX)^{-1}a$ has to be computed. For the design actually used this is 6.4×10^{-7}. Now δ is 2. The power of a two-sided t test at level 0.05 and with $18 - 4 = 14$ degrees of freedom is 0.46 (from table B 5 page 1346).

Take notice that you need to specify α, β_3/σ (or even β_3 and σ) and the design!

SAMPLE SIZE NEEDED using t test: SAND and FIBRE example.

Now for the same assumed values of the parameters how many replicates of the basic design (using 9 combinations of sand and fibre contents) would I need to get a power of 0.95? The matrix X^TX for m replicates of the design actually used is m times the same matrix for 1 replicate. This means that $a^T(X^TX)^{-1}a$ will be $1/m$ times the same quantity for 1 replicate. Thus the value of δ for m replicates will be $\sqrt{m}$ times the value for our design, which was 2. With m replicates the degrees of freedom for the t-test will be $18m - 4$. We now need to find a value of m so that in the row in Table B 5 across from $18m - 4$ degrees of freedom and the column corresponding to

$$\delta = 2\sqrt{m}$$

we find 0.95. To simplify we try just assuming that the solution m is quite large and use the last line of the table. We get δ between 3 and 4 – say about 3.75. Now set $2\sqrt{m} = 3.7$ and solve to find $m = 3.42$ which would have to be rounded to 4 meaning a total sample size of $4 \times 18 = 72$. For this value of m the non-centrality parameter is actually 4 (not the target of 3.75 because of rounding) and the power is 0.98. Notice that for this value of m the degrees of freedom for error is 66 which is so far down the table that the powers are not much different from the ∞ line.

POWER of F test: SAND and FIBRE example.

Now consider the power of the test that all the higher order terms are 0 in the model

$$Y_i = \beta_0 + \beta_1 S_i + \beta_2 F_i + \beta_3 F_i^2 + \beta_4 S_i^2 + \beta_5 S_i F_i + \epsilon_i$$

that is the power of the F test of $\beta_3 = \beta_4 = \beta_5 = 0$.

You will need to specify the non-centrality parameter for this F test. In general the noncentrality parameter for a F test based on ν_1 numerator degrees of freedom is given by

$$E(\text{Extra SS})/\sigma^2 - \nu_1 \, .$$

This quantity needs to be worked out algebraically for each separate case, however, some general points can be made.

- Write the full model as

$$Y = X_1\beta_1 + X_2\beta_2 + \epsilon$$

  and the reduced model as

$$Y = X_1\beta_1 + \epsilon$$

- The Extra SS is the difference between two Error sums of squares. One is for the full model and

$$E(\text{ErrorSS}_{\text{FULL}}) = \text{ErrorDF}_{\text{FULL}}\sigma^2$$

  because we assume that the FULL model is correct.

- The Error SS for the reduced model is

$$Y^T(I - H_1)Y$$

  where $H_1 = X_1(X_1^TX_1)^{-1}X_1^T$. Replace Y by its formula $X_1\beta_1+X_2\beta_2+\epsilon$ from the full model equation and take expected value. The answer is

$$\sigma^2[(n - p_1) + \beta_2^TX_2^T(I - H_1)X_2\beta_2]$$

  where $p_1 =$ is the rank of X_1. This makes the non-centrality parameter $\beta_2^TX_2^T(I - H_1)X_2\beta_2]$.