```
function [G_FilterNumber] = IP_GFilter(Groups,num_g,nk,G_thres)

    % Groups is a vector of class indices associated with each observation in Peptide_Vector. This vector
        % is numeric and sequential, which means that if there are 4 classes, then each position in
        % Peptide_Vector will be represented in "Groups" as a 1,2,3 or a 4.  If the data are in another
        % format, such as a Cell structure it would need to be converted.

    % num_g is a constant that defines how many groups are present in the data.

    % nk is a vector of length num_g that identifies how many samples are associated with group k

    % G_thres is the significance value required to identify the minimum number of observations required
      % in a single group (G-filter): (0.05 in paper)

    G_FilterNumber = zeros(1,num_g);
    N = length(Groups);
    for i = 1:num_g
        Peptide_Vector = NaN(N,1);
        temp = NaN(nk(i),1);
        GTest_pvalue = 1;
        t = 0;
        % if GTest_pvalue is less than the defined threshold than stop, otherwise t must be increased.  If t is equal to nk(i)
        % and GTest_pvalue is still not less than threshold than set equal to the size of the group.
    while GTest_pvalue > G_threshold && t < nk(i)
            t = t+1;
            temp(t) = 1;
            Peptide_Vector(Groups == i) = temp;
            GTest_pvalue = GTest(Peptide_Vector,Groups);
        end
        G_FilterNumber(i) = t;
    end
end
```

Figure S5. Generic MatLab® code for computing the number of peptides retained in each group using the G-Filter as described in Equation 3.