# Constrast and Classical Inference

## Lecture Notes

Ahmet Ademoglu, *PhD*
Bogazici University
Institute of Biomedical Engineering

Some concepts and illustrations in this lecture are adapted from the textbook,
**Statistical Parametric Mapping: The Analysis of Functional Brain Images**, Editors: K. Friston, J. Ashburner, S. Kiebel, T. Nichols and W. Penny, *Academic Press*, 2006.

# t-test in GLM

General Linear Model:

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \mathbf{e}$$

with $\mathbf{e} \sim \mathcal{N}(\mathbf{0}, \mathbf{V})$.
General Least Squares Solution:

$$\mathbf{b} = \mathbf{X}^{-}\mathbf{y}$$

where $\mathbf{X}^{-} = (\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{V}^{-1}$ and is the pseudo inverse of $\mathbf{X}$.
The contrast has a distribution $\hat{\mathbf{c}}^T\mathbf{b} \sim \mathcal{N}(\mathbf{c}^T\mathbf{b}, (\mathbf{X}^T\mathbf{V}^{-\mathbf{1}}\mathbf{X})^{-1})$
and the $t-$ statistics : $t_\nu = \frac{\mathbf{c}^T\hat{\mathbf{b}}}{\sqrt{\mathbf{c}^T(\mathbf{X}^T\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{c}}}$
Effective degrees of freedom $\nu = \frac{trace(\mathbf{RV})^2}{trace(\mathbf{RVRV})}$

# F-Test

$$\mathbf{X} = \begin{bmatrix} X_1 & \vdots & X_2 \end{bmatrix} \qquad \beta = [\beta_{\mathbf{1}} \vdots \beta_{\mathbf{2}}]^T,$$

$$\mathbf{Y} = \begin{bmatrix} X_1 & \vdots & X_2 \end{bmatrix} \begin{bmatrix} \beta_{\mathbf{1}} \\ \dots \\ \beta_{\mathbf{2}} \end{bmatrix} + \mathbf{e}$$

$\mathcal{H} : \beta_1 = 0$

Residual Sum of Squares $S(\beta) = (\mathbf{RY})^T\mathbf{RY}$

$\mathbf{R} = \mathbf{I} - \mathbf{X}\mathbf{X}^-$

$\mathbf{R_0} = \mathbf{I} - \mathbf{X_2}\mathbf{X_2}^-$

$\mathbf{M} = \mathbf{R_0} - \mathbf{R}$

$$F = \frac{S(\beta_2) - S(\beta)}{S(\beta)} \frac{J-p}{p-p_2} = \frac{\mathbf{Y}^T\mathbf{MY}}{\mathbf{Y}^T\mathbf{RY}} \frac{J-p}{p-p2} \sim F_{p_1, J-p}$$

# F-Test for the DCT coefficients modeling the drift

```
N=100;xd=[0:N-1]'; % Drift Signal
w = ones(1,N)/sqrt(N); % DC component
d=sqrt (2/N)*cos([1:N-1]'*[1:2:2*N-1]*pi/(2*N)); % higher frequency components
d=[w;d]; y = dct(xd); % estimating the DCT coefficients
k = 6; % Number of DCT components to model the drift
[hrf,p] = spm_hrf(1);
x= zeros(N,1); x([11:20 41:60 71:90]) = 1;
y = convmtx (x,length(hrf))*hrf;
y = y(1:length(x),1:end);
y1 = y + randn(size(y))*0.4 ; % Add white noise to BOLD with sigma = 0.4
Y = y1 + 0.04*xd; % Add a drift to BOLD
% Design Matrix
X = [ y-mean(y)  d(1:k,:)' ] ;
beta = X\Y;
Xpinv = pinv(X);
R = eye(length(Y),length(Y)) - X*Xpinv;
X0 = X(:,1) ;
X0pinv = pinv(X0);
R0 = eye(length(Y),length(Y)) - X0*X0pinv;
r0 = R0*Y; % S(beta_2)  reduced model Sum of squares
r = R * Y;   % S(beta)  Full model Sum of squares
sigma = std(r);
V = sigma^2*eye(length(y1),length(y1)); % covariance of error
nu0 = (trace ( (R - R0)*V))^2/trace((R-R0)*V*(R-R0)*V); % df for reduced model
nu = (trace (R*V))^2/trace(R*V*R*V); % df for full model
F = ( ( r0'*r0 - r'*r )*nu )/ (nu0 * (r'*r ) ) % F(nu,nu0) distribution
1-fcdf(F,nu0,nu)  % p value
```

This formulation of the $F-$statistic has an advantage that $\mathbf{X}$ can be partitioned into any two sets of linear combinations of the regressors.
A contrast matrix $\mathbf{C}^T$ is used to specify a subspace of the design matrix $\mathbf{X}_C = \mathbf{X}\mathbf{C}$.
The orthogonal contrast to $\mathbf{C}$ is $\mathbf{C}_0 = \mathbf{I} - \mathbf{C}\mathbf{C}^-$.
$\mathbf{X}_0 = \mathbf{X}\mathbf{C}_0$ is the design matrix of the reduced model.
If we wish to test the effects $\mathbf{X}_C$ can explain, after fitting the reduced model $\mathbf{X}_0$ we can use a matrix that projects the data onto the subspace of $\mathbf{X}_C$, which is orthogonal to $\mathbf{X}_0$.
This subspace can be denoted by $\mathbf{X}a$.
The projection operator $\mathbf{M}$ to this subspace $\mathbf{X}a$ is $\mathbf{M} = \mathbf{R}_0 - \mathbf{R}$ where $\mathbf{R}_0 = \mathbf{I}_J - \mathbf{X}_0\mathbf{X}_0^-$ and $\mathbf{R} = \mathbf{I}_J - \mathbf{X}\mathbf{X}^-$ are the residual forming matrices of reduced and full models, respectively.

## The $F-$statistics

$$F = \frac{\mathbf{Y}^T\mathbf{M}\mathbf{Y}}{\mathbf{Y}^T\mathbf{R}\mathbf{Y}} \frac{J-p}{p-p2} \sim F_{p_1, J-p}$$

where $p_1$ is the rank of $\mathbf{X}_a$.
Since $\mathbf{M}$ projects $\mathbf{Y}$ onto a subspace within $\mathbf{X}$,

$$F = \frac{\hat{\beta}^T\mathbf{X}^T\mathbf{M}\mathbf{X}\hat{\beta}}{\mathbf{Y}^T\mathbf{R}\mathbf{Y}} \frac{J-p}{p1} \sim F_{p_1, J-p}$$

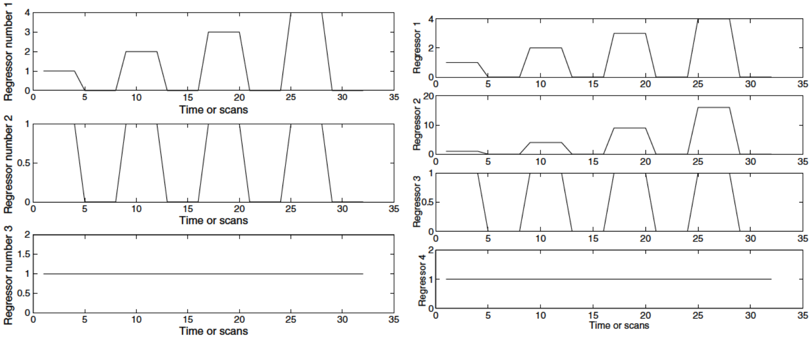
This means we can compute an $F-$statistic conveniently for any user-specified contrast without a re-parameterization.

# F-Test for the DCT coefficients modeling the drift with F-Contrast

```
clear all;
N=100;xd=[0:N-1]'; % Drift Signal
w = ones(1,N)/sqrt(N); % DC component
d=sqrt (2/N)*cos([1:N-1]'*[1:2:2*N-1]*pi/(2*N)); % higher frequency components
d=[w;d]; y = dct(xd); % estimating the DCT coefficients
k = 6; % Number of DCT components to model the drift
[hrf,p] = spm_hrf(1);
x= zeros(N,1); x([11:20 41:60 71:90]) = 1;
y = convmtx (x,length(hrf))*hrf;
y = y(1:length(x),1:end);
y1 = y + randn(size(y))*0.4 ; % Add white noise to BOLD with sigma = 0.4
Y = y1 + 0.04*xd; % Add a drift to BOLD
% Design Matrix
X = [ y-mean(y)  d(1:k,:)' ] ;
beta = X\Y;
C = [zeros(k,1) diag(ones(k,1))]';
XC = X*C;
C_ = pinv(C);
C0 = eye(size(C,1)) - C*C_;
X0 = X*C0; % reduced model
R0 = eye(N)  - X0*pinv(X0); % residual froming matrix for reduced model
R  =  eye(N)  - X*pinv(X); % residual froming matrix for full model
M = R0 - R;
nu = rank(R)
nu0 = rank(M)
F  = (beta'*X'*M*X*beta*nu)/(Y'*R*Y*nu0)
1-fcdf(F,nu0,nu)  % p value
```

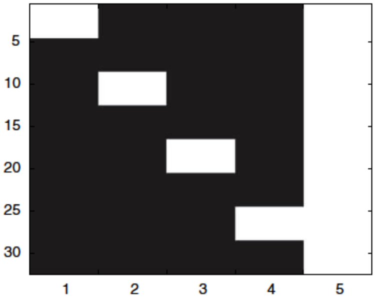
# Constructing Models

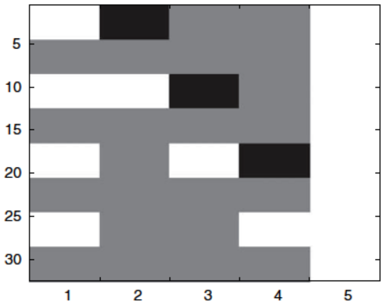
**fMRI experiment:**
Subject presses a device with four different force levels.



Different force levels modeled with different covariates

Main effect of force : 1st regressor
Interactions : regressors 2, 3 and 4

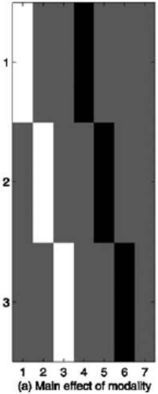



- Is there an overall difference between force levels and the rest condition?
- Are there any differences between force levels?

# Constructing F-Contrasts

Words presented visually (V) or orally (O) with 3 categories C1-C2-C3



# Types of ANOVA

The ANOVA uses $F-$tests to examine the main effects and interactions.

| Factors | Levels | Simple | Repeated measures |
|---|---|---|---|
| 1 | 2 | Two-sample *t*-test | Paired *t*-test |
| 1 | K | One-way ANOVA | One-way ANOVA within-subject |
| M | $K_1, K_2, .., K_M$ | M-way ANOVA | M-way ANOVA within-subject |

# One way Between Subject ANOVA

$$y_n = \tau_k + \mu + e_n$$

$\tau_k$ : treatment effects
$k$: index for group assignment
$\mu$: grand mean or constant term
If the factor is significant, it is significantly better than the simple model

$$y_n = \mu + e_n$$

# Example: $K = 4$ Groups, $N = 12$ subjects

$[\tau_1 \tau_2 \tau_3 \tau_4 \mu]^T$

$y_n = \tau_k + \mu + e_n$

versus

$y_n = \mu + e_n$

$\mathcal{H}_0$: $\tau_1 = \tau_2 = \tau_3 = \tau_4$

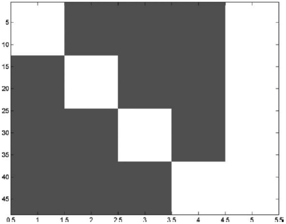
$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$

$\mathbf{X} = [\mathbf{I}_K \otimes \mathbf{1}_N \;\; \mathbf{1}_{N \cdot K}]$

Effects of interest

$F-$ Contrast:

$$\mathbf{C}^T = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$



# One way within Subject ANOVA

$y_{nk} = \tau_k + \pi_n + \mu + e_{nk}$
versus
$y_{nk} = \pi_n + e_{nk}$
$\tau$ : within subject $K$ treatment effects
$\pi$ : Subjects effects

# Example: $K = 4$ and $N = 12$

$[\tau_1 \tau_2 \tau_3 \tau_4 \pi_1 \pi_2 \ldots \pi_{12}]$

$y_{nk} = \tau_k + \pi_n + e_{nk}$

versus

$y_{nk} = \pi_n + e_{nk}$

$\mathcal{H}_0$: $\tau_1 = \tau_2 = \tau_3 = \tau_4$

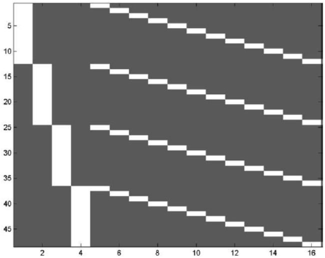
$\mathbf{y} = \mathbf{X}\beta + \mathbf{e}$

$\mathbf{X} = [\mathbf{I}_K \otimes \mathbf{1}_N \;\; \mathbf{1}_K \otimes \mathbf{I}_N]$

Effects of interest

$F-$ Contrast:

$$\mathbf{C}^T = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & 1 & -1 & 0 & 0 & \ldots & 0 \\ 0 & 0 & 1 & -1 & 0 & \ldots & 0 \end{bmatrix}$$



# 2 way within Subject ANOVA $K_1$ by $K_2$

$y_{nkl} = \tau_{kl} + \pi_n + e_{nkl}$
versus
$y_{nkl} = \pi_n + e_{nk}$
$k$ and $l$ index the levels of factors $A$ and $B$
$\tau$ : within subject treatment effects
$\pi$ : Subjects effects

# 2-way within Subject ANOVA $K_1 = 2$, $K_2 = 2$

$$P = K_1 \times K_2 = 4 \qquad N = 12 \qquad J = NP = 48$$

| | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| | $A_1B_1$ | $A_1B_2$ | $A_2B_1$ | $A_2B_2$ |
| Main Effect of A | $[\,1$ | $1$ | $-1$ | $-1\,]$ |
| Main Effect of B | $[\,1$ | $-1$ | $1$ | $-1\,]$ |
| Interaction, $A \times B$ | $[\,1$ | $-1$ | $-1$ | $1\,]$ |

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e} \qquad \mathbf{X} = [\mathbf{I}_P \otimes \mathbf{1}_N \; \mathbf{1}_P \otimes \mathbf{I}_N]$$

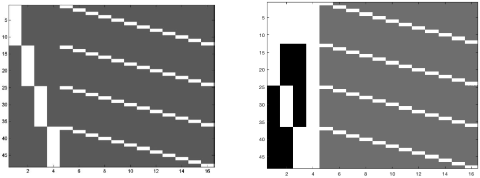
$\beta = [\tau_{kl} \; \pi_n]$ 2$-$ factor ANOVA : 2 main effects and 1 interaction

$$\mathbf{y} = \mathbf{X}\beta + \mathbf{e} = \mathbf{X}\tilde{\mathbf{C}}^{-T}\tilde{\mathbf{C}}^{T}\beta + \mathbf{e} = \mathbf{X}_r\tilde{\beta} + \mathbf{e}$$

$$\tilde{\mathbf{C}}^{-T} = \begin{bmatrix} \mathbf{C}^T & \mathbf{0}_{4,12} \\ \mathbf{0}_{12,4} & \mathbf{I}_{12} \end{bmatrix} \quad \mathbf{C}^T = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \end{bmatrix}$$

```
CT = [- [-1 -1 1 1 ; -1 1 -1 1 ]; 1 -1 -1 1 ; 1 1 1 1 ];
X=[kron(eye(4,4),ones(12,1)) kron(ones(4,1),eye(12,12) )];
CT_til_1 = blkdiag(inv(CT)*4,eye(12,12));
imagesc(X);colormap gray;
imagesc(X*CT_til_1);colormap gray;
```

$\beta = [\tau_{11}\tau_{12}\tau_{21}\tau_{22}\ \pi_1 \ldots \pi_{12}]$ $\tilde{\beta} = [\tau_1^A \tau_1^B \tau_1^{AB} m]$



# 2-way within Subject ANOVA with $K_1 \times K_2$ and $N$ Subjects

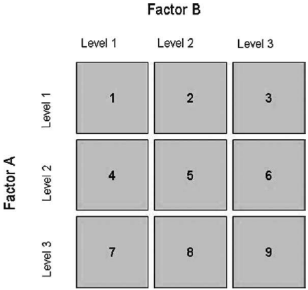
$$A_1B_1 \;\; A_1B_2 \ldots A_1B_{K_2} | A_2B_1 \;\; A_2B_2 \ldots A_2B_{K_2} | \ldots | A_{K_1}B_1 \;\; A_{K_1}B_2 \ldots A_{K_1}B_{K_2}$$

- $\tau_q^A$: differences between each successive level $q = 1 \ldots K_1 - 1$ of factor $A$ averaged over factor $B$.
- $\tau_r^B$: differences between each successive level $r = 1 \ldots K_2 - 1$ of factor $B$ averaged over factor $A$.
- $\tau_{qr}^{AB}$:differences between differences of each level $q = 1 \ldots K_1 - 1$ of factor $A$ across each level of $r = 1 \ldots K_2 - 1$ of factor $B$.
- $m$: mean treatment effect
- $\pi_n$ :subject effect

$$\tilde{\beta} = [\tau_1^A \ldots \tau_q^A \; \tau_1^B \ldots \tau_r^B \; \tau_{11}^{AB} \ldots \tau_{qr}^{AB} \; m \; \pi_1 \ldots \pi_N]$$

# Pooled Error Model for a 3 × 3 Repeated Measures ANOVA

| 1 | 2 | … | $K_3$ | … | $P$ |
|---|---|---|---|---|---|
| $A_1 B_1 C_1$ | $A_1 B_1 C_2$ | … | $A_1 B_1 C_{K_3}$ | … | $A_{K_1} B_{K_2} C_{K_3}$ |




For a 3 × 3 ANOVA:

$\mathbf{C}_1 = \mathbf{C}_2 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$

$\mathbf{D}_1 = \mathbf{D}_2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}^T$

Main effect of $A$: $\mathbf{D}_1 \otimes \mathbf{C}_2$

Main effect of $B$: $\mathbf{C}_1 \otimes \mathbf{D}_2$

Interaction : $\mathbf{D}_1 \otimes \mathbf{D}_2$

$\tilde{\beta} = [\tau_1^A \; \tau_2^A \; \tau_1^B \; \tau_2^B \; \tau_{11}^{AB} \tau_{12}^{AB} \; \tau_{21}^{AB} \; \tau_{22}^{AB} \; m \; \pi_n]$

$\mathcal{H}_0$ for the main effect of A :

$\tau_1^A = \tau_2^A = 0.$

$\mathcal{H}_0$ for the interaction : $\tau_{11}^{AB} = \tau_{12}^{AB} = \tau_{21}^{AB} = \tau_{22}^{AB} = 0.$

# $F-$ Contrasts of the Main Effects and Interactions

$C_m = \mathbf{1}_{K_m} \qquad D_m = -orth(diff(\mathbf{I}_{K_m})^T)$

$diff(A)$ : Column differences of matrix $A$

$\mathbf{C}_1 = \mathbf{C}_2 = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}^T$

$$\mathbf{D}_1 = \mathbf{D}_2 = \begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix}^T$$

$$D_1 \otimes C_2 = \begin{bmatrix} 1 & 1 & 1 & -1 & -1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & -1 & -1 & -1 \end{bmatrix}^T$$

$$C1 \otimes D_2 = \begin{bmatrix} 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 & 0 \\ 0 & 1 & -1 & 0 & 1 & -1 & 0 & 1 & -1 \end{bmatrix}^T$$

$$D1 \otimes D_2 = \begin{bmatrix} 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 & -1 & 1 & 0 \end{bmatrix}^T$$

## $F-$Contrast Matrices for 3 by 3 ANOVA

```
C1 = [1 1  1]'
C2 = [1 1  1]'
D1 = -diff(eye(3),[],2)
D2 = -diff(eye(3),[],2)
kron(D1,C2)
kron(C1,D2)
kron(D1,D2)
```

# Exercise

There are three models of cars (columns) and two factories (rows). The reason there are six rows in mileage instead of two is that each factory provides three cars of each model for the study. The data from the first factory is in the first three rows, and the data from the second factory is in the last three rows.

| | *Car Models* | | |
|---|---|---|---|
| | 33.3000 | 34.5000 | 37.4000 |
| | 33.4000 | 34.8000 | 36.8000 |
| *Mileage* | 32.9000 | 33.8000 | 37.6000 |
| | 32.6000 | 33.4000 | 36.6000 |
| | 32.5000 | 33.7000 | 37.0000 |
| | 33.0000 | 33.9000 | 36.7000 |

Determine the effect of car model and factory on the mileage rating of cars using 2-way ANOVA based on GLM and F-contrasts *i.e.* main effect of factory, main effect of car model and interaction between factory and car model.

# 2-Way ANOVA for car model and factory

```
load mileage
N = 3; % number of subjects
K1 = 2;
K2 = 3;
J =  N*K1*K2;
% order the data as
Y = [ reshape(mileage(1:N,:),N*K2,1) ;  reshape(mileage(N+[1:N],:),N*K2,1) ];
X =[ kron(eye(K1*K2,K1*K2),ones(N,1))    ] ;
imagesc(X);colormap gray;pause ;close
C1 = ones(K1,1);
C2 = ones (K2,1);
D1 = - (diff(eye(K1,K1),[],2));
D2 = - (diff(eye(K2,K2),[],2));
X1 = kron(D1,C2)';
X2 = kron(C1,D2)';
X3 = kron(D1,D2)';
CT  = [ X1' X2' X3' ones(K1*K2,1)  ];
X = X* CT; % rotating the design matrix
imagesc(X);colormap gray;pause ;close
beta = X\Y; % estimated parameter vector

% Full model
R  =  eye(J)  - X*pinv(X); % residual froming matrix for full model
p  = rank(X);
nu = J - p;
```

# 2-Way ANOVA for car model and factory

```
% Main effect of Factory
C =  zeros(1,size(X,2))';
C(1) = 1;
XC = X*C;
C_ = pinv(C);
C0 = eye(size(C,1)) - C*C_;
X0 = X*C0;
R0 = eye(J)  - X0*pinv(X0);
M = R0 - R;
p2 = rank(X0);
nu0   = p-p2
F(1)  = (beta'*X'*M*X*beta*nu)/(Y'*R*Y*nu0)
pval(1)  = 1-fcdf(F(1),nu0,nu)  % p value

% Main effect of Model
C =  zeros(2,size(X,2))';
C(2,1) = 1;
C(3,2) = 1;
XC = X*C;
C_ = pinv(C);
C0 = eye(size(C,1)) - C*C_;
X0 = X*C0;
R0 = eye(J)  - X0*pinv(X0);
M = R0 - R;
p2 = rank(X0);
nu0   = p-p2
F(2)  = (beta'*X'*M*X*beta*nu)/(Y'*R*Y*nu0)
pval(2)  = 1-fcdf(F(2),nu0,nu)  % p value
```

## 2-Way ANOVA for car model and factory

```
%   Effect of Factory X Model interaction
C =  zeros(2,size(X,2))';
C(4,1) = 1;
C(5,2) = 1;
XC = X*C;
C_ = pinv(C);
C0 = eye(size(C,1)) - C*C_;
X0 = X*C0;
R0 = eye(J)  - X0*pinv(X0);
M = R0 - R;
p2 = rank(X0);
nu0   = p-p2
F(3)  = (beta'*X'*M*X*beta*nu)/(Y'*R*Y*nu0)
pval(3)  = 1-fcdf(F(3),nu0,nu)  % p value

%*************************************
% MATLAB STATISTICAL TOOLBOX
cars = 3;
[p,tbl,stats] = anova2(mileage,cars);
```