$y = x\beta + \epsilon$
$\hat{\beta} = (X^{'}X)^{-1}X^{'}y$
$\hat{y} = X\hat{\beta} = Hy$
$H = X(X^{'}X)^{-1}X^{'} \quad Hy = \hat{y}$
$M = C(X^{'}X)^{-1}C^{'}$
$\hat{\epsilon} = y - \hat{y} = y - Hy = (1 - Hy)$
$\hat{\theta} = C\hat{\beta}$
$\hat{Var}(\hat{\beta}) = \hat{\sigma^2}(X^{'}X)^{-1}$
$\hat{Var}(\hat{\theta}) = \hat{\sigma^2}(C(X^{'}X)^{-1}C^{'})$
$SSH = (\hat{\theta} - \theta_0)^{'}M^{-1}(\hat{\theta} - \theta_0)$
$SSE = (y - \hat{y})^{'}(y - \hat{y}) = y^{'}[I - H]y = \epsilon'\epsilon$
Use MSE for $\hat{\sigma^2}$, use diag
$se = \sqrt{\hat{Var}} \quad MS* = SS*/DF*$
$F = MST/MSE \quad df_{total} = n - 1$
t-value= $\hat{\beta}/se$ p-value= $F(df_{model}, df_{error})$
95% CI $\hat{\theta} \pm 1.96\sqrt{MSE(M)}$
95% PI $\hat{y_h} \pm t(\alpha/2, n-r)se(\hat{y_h})$
$F_{obs} = \dfrac{(\hat{\theta} - \theta_0)^{'}M^{-1}(\hat{\theta} - \theta_0)/a}{\hat{\sigma^2}} \; a = rank(C)$

Matrix
$(X_{n\times p})^{'}(X_{n\times p})\hat{\beta} = (X_{n\times p})^{'}y_{n\times 1}$
$(AB)^{'} = B^{'}A^{'} \quad A^{-1}A = AA^{-1} = I$
$(AB)^{-1} = B^{-1}A^{-1} \quad (A^{'})^{-1} = (A^{-1})^{'}$
$|A^{-1}| = \dfrac{1}{|A|} \quad \dfrac{1}{|A|}\begin{vmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{vmatrix}$
det(inv)=inv(det) evector $Ax = \lambda x$ e val=λ
char eqn ($|A - \lambda I| = 0$) $= det\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix}$
full rank $\Leftrightarrow$ no $\lambda = 0$ det=0 $\Leftrightarrow \geq 1$ $\lambda = 0$
X full rank for $\hat{\beta}$ unique
E(rand mat)=mat of E
$Cov(Y) = E[(Y - \mu)(Y - \mu)^{'}] = \mathbf{\Sigma}$
$\begin{vmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \cdots & \cdots & \sigma_{nn} \end{vmatrix}$
where $\sigma_{ij} = E[(Y_i - \mu_i)(Y_j - \mu_j)^{'}]$
$Cov(AY + b) = ACov(Y)A^{'} = A\Sigma A^{'}$
$Cov(W, Y) = E[(W - \gamma)(Y - \mu)^{'}]$
MVN Σ pos def
1) **lin trans** of mvn yields mvn.
If $X \sim N_n(\mu, \Sigma)$ and $Y = AX + b$
with $A_{r\times n}$ matrix of constants and $\boldsymbol{b}_{r\times 1}$ vector of constants.
Then $Y \sim N_r(A\mu + b, A\Sigma A^{'})$
2) **lin combin** of ind mvns is mvn
$X_1, \ldots, X_k$ iid $X_i \sim N_n(\mu_i, \Sigma_i)$
Let $Y = a_1X_1 + \cdots + a_kX_k$
Then $Y \sim N(\mu^*, \Sigma^*)$ where $\mu^* = \sum_{i=1}^{k} a_i\mu_i$
and $\Sigma^* = \sum_{i=1}^{k} a_i^2\Sigma_i$
3) **Marginals** of mvn are mvn
If $X \sim N_n(\mu, \Sigma)$
Partition: $X = \binom{X_1}{X_2}$ where X_1 is $r \times 1$ and X_2 is $(n-r) \times 1$
$\mu = \binom{\mu_1}{\mu_2}$ where μ_1 is $r \times 1$ and μ_2 is $(n-r) \times 1$
$\Sigma = \begin{vmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{vmatrix}$
where Σ_{11} is $r \times r$, Σ_{21} is $(n-r) \times r$ and Σ_{22} is $(n-r) \times (n-r)$
Marginals: $X_1 \sim N_r(\mu_1, \Sigma_{11})$
$X_2 \sim N_{(n-r)}(\mu_2, \Sigma_{22})$
4) **Conditionals** of mvn are mvn.
Suppose $X \sim N_n(\mu, \Sigma)$
Partition same way
$X_1|X_2 = x_2 \sim N_r(\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2), \Sigma^*)$
where $\Sigma^* = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$

Estimable, Testable
If X is full rank, any (nonzero) $\boldsymbol{C}$ gives estimable $\boldsymbol{\theta}$
$rank(X) = rank(X'X) =$ number of estimable parameters, and for $\hat{\boldsymbol{\theta}} = \boldsymbol{C\beta}$, check estimability.
If X full rank then $\boldsymbol{\theta}$ is testable $\Leftrightarrow$
$\boldsymbol{C}$ is full rank a or $\boldsymbol{M}$ is full rank a (because any $\boldsymbol{\theta}$ is estimable)

T-tests
One-sided Test (only for scalar hypothesis)
$F_{obs}(1, n-r) = tval^2(n-r)$
$H_A : \theta < \theta_0$ Uses α $H_A : \theta > \theta_0$ uses $(1 - \alpha)$
Two-sided $H_A : \theta \neq \theta_0$ uses $\alpha/2$ and $(1 - \alpha/2)$
In all cases reject H_0 if |test stat| > |crit value|

F-tests
Reject H_0 if $f_{obs} > f_{crit}$
Two-sided uses α crit value
One-sided $\theta < \theta_0$ uses 2α and requires appropriate sign of difference
All linear model GLH tests correspond to comparing two models, the "full" model, $\boldsymbol{y = X\beta + \epsilon}$ and a reduced model defined by constraints
For a single coefficient β_j we can test $H_0 : \beta_j = 0$ if β_j is estimable
$t = \frac{\hat{\beta}_j - 0}{\sqrt{var(\hat{\beta}_j)}}$
Obtain estimate $\hat{\sigma}^2 = MSE$
if know σ^2 then $t \sim N(0,1)$
if estimate σ^2 from data then $t \sim t_{dfE}$

Sum of Squares Decomp
$USS(total) = y^{'}y = \sum_{i=1}^{n} y_i^2$
$USS(model) = y^{'}X(X^{'}X)^{-1}X^{'}y$
$= y^{'}Hy = \sum_{i=1}^{n}\hat{y}_i^2$
$USS(total) = USS(model) + SSE$
$CSS(total) = USS(total) - SSI$
$= \sum_{i=1}^{n}(y_i - \bar{y})^2$
$CSS(total) = CSS(model) + SSE$

R squared,Corr
$R^2_{adj} = 1 - \frac{SSE/(n-r)}{CSS(Total)/(n-1)}$
adjusted for the df. It will only increase on adding a variable to the model if the variable reduces the mean square for error
$\rho = Corr(X, Y) = \frac{Cov(x,y)}{\sqrt{var(x)var(y)}}$
estimate ρ using pearsons coef of corr
$R = \frac{\sum_{i=1}^{n}(X_i - \bar{X})Y_i - \bar{Y}}{\sqrt{(\sum_{i=1}^{n}(X_i - \bar{X})^2)(\sum_{i=1}^{n}(Y_i - \bar{Y})^2)}}$
$(X^{'}X)^{-1} = \begin{vmatrix} .308 & -.06 & -.017 \\ -.06 & -.025 & -.004 \\ -.017 & -.004 & .006 \end{vmatrix} \; X^{'}y = \begin{vmatrix} 405 \\ 1402 \\ 2350 \end{vmatrix}$

Source	DF	SS	MS	Fval	$P > F$
Model	2	79	239.5	349.6	$< .001$
Error	97	11	.113		
Ctotal	99	90			

param	est	se	tval	$p > \|t\|$
x0	.67	.187	3.58	.009
x1	1.35	.053	25.47	$< .001$
x2	1.61	.026	61.81	$< .001$

2) c) Test $H_0 : \beta_1 = 1$
t-test=$\frac{\hat{\beta}_1 - 1}{se(\hat{\beta}_1)} = 6.6 \sim t_{97}$ $6.6 > 1.96$ reject H_0
d) Test $H_0 : \beta_1 = \beta_2 = 1$
$B_1 = 1 \quad B_2 = 1$
$C = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} \; \theta_0 = \begin{vmatrix} 1 \\ 1 \end{vmatrix} \; H_0 : C\theta = \theta_0 \; \hat{\theta} = \begin{vmatrix} 1.35 \\ 1.61 \end{vmatrix}$
calc M calc M^{-1}
F-test=$\frac{43.83}{.113} = 387.8 \sim F_{2,97}$
e) 95% CI of $\beta_1 + \beta_2$ (θ)
$\hat{\theta} = \hat{\beta}_1 + \hat{\beta}_2 = 2.96$ $C = [0, 1, 1]$
$se(\hat{\theta}) = \hat{\sigma}^2_{\hat{\theta}} = \sqrt{.0026}$
$2.96 \pm 1.96\sqrt{.0026} = [2.86, 3.06]$
f) Transform x_1, x_2 to $z_1 = x_1 - 2$, $z_2 = x_2 - 4$
Refit with $y^* = \beta_0^* + \beta_1^*z_1 + \beta_2^*z_2$
$= (\beta_0^* - 2\beta_1^* - 4\beta_2^*) + \beta_1^*x_1 + \beta_2^*x_2$
$\beta_0^* = \beta_0^* + 2\beta_1^* + 4\beta_2^* = 9.8$ $C = [1, 2, 4]$
$\sigma^2 = \hat{\sigma}^2C(X^{'}X)^{-1}C^{-1} = .085$ $t = 115.3$
3) f) is $H_0 : \beta_0 - \beta_2 = 0$ and $\beta_1 + 2\beta_2 = 2$ and $2\beta_0 + \beta_1 = 2$ testable? Reduce to ETH?
$y = \begin{vmatrix} 0 \\ 2 \\ 3 \\ 6 \\ 10 \end{vmatrix} \; X = \begin{vmatrix} 1 & 6 & 11 \\ 1 & 7 & 13 \\ 1 & 8 & 15 \\ 1 & 9 & 17 \\ 1 & 11 & 21 \end{vmatrix} \; \beta = \begin{vmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{vmatrix}$
$C = \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 2 & 1 & 0 \end{vmatrix}$ reduces $\begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{vmatrix}$ 2pivots $\theta_0 = \begin{vmatrix} 0 \\ 2 \\ 2 \end{vmatrix}$
not FR, not testable, reduces to C^* testable
$C^* = \begin{vmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{vmatrix} \; \theta_0^* = \begin{vmatrix} 0 \\ 2 \end{vmatrix} \; H_0 : \begin{vmatrix} \beta_0 - \beta_2 = 0 \\ \beta_1 + 2\beta_2 = 2 \end{vmatrix}$
1) a) $x = \begin{vmatrix} x_1 \\ x_2 \\ x_3 \end{vmatrix} \sim N(0, \Sigma)$ $\Sigma = \begin{vmatrix} 1 & 0 & .6 \\ 0 & 1 & .5 \\ .6 & .5 & 1 \end{vmatrix}$
$c = \begin{vmatrix} 2 & 1 & -1 \end{vmatrix}$
$2x_1 + x_2 - x_3 \sim N(c\mu, c\Sigma c') = N(0, 2.6)$
b) $Cov(x_1 - x_2, 2x_2 + x_3) = c_1\Sigma c_2' = -1.9$
$c_1 = \begin{vmatrix} 1 & -1 & 0 \end{vmatrix}$ $c_2 = \begin{vmatrix} 0 & 2 & 1 \end{vmatrix}$

Quantity	Definition	Description
$\overline{x}$	$\sum x_i/n$	Sample average of x
$\overline{y}$	$\sum y_i/n$	Sample average of y
SXX	$\sum(x_i - \overline{x})^2 = \sum(x_i - \overline{x})x_i$	Sum of squares for the xs
SD_x^2	SXX$/(n-1)$	Sample variance of the xs
SD_x	$\sqrt{SXX/(n-1)}$	Sample standard deviation of the xs
SYY	$\sum(y_i - \overline{y})^2 = \sum(y_i - \overline{y})y_i$	Sum of squares for the ys
SD_y^2	SYY$/(n-1)$	Sample variance of the ys
SD_y	$\sqrt{SYY/(n-1)}$	Sample standard deviation of the ys
SXY	$\sum(x_i - \overline{x})(y_i - \overline{y}) = \sum(x_i - \overline{x})y_i$	Sum of cross-products
s_{xy}	SXY$/(n-1)$	Sample covariance
r_{xy}	$s_{xy}/(SD_xSD_y)$	Sample correlation

[a]In each equation, the symbol Σ means to add over all n values or pairs of values in the data.

Source of Variation	df	SS	MS	F_{obs}	p
Intercept	1	SSI	SSI	SSI/MSE	
Model(uncorrected)	q	USS(model)	USSM/q	(USSM/q)/MSE	
Model(corrected)	q-1	CSS(model)	CSSM/q-1	(CSSM/q-1)/MSE	
Error(residual)	N-q	SSE	SSE/(N-q)=MSE	-	-
Total(uncorrected)	N	USS(total)	-	-	-
Total(corrected)	N-1	CSS(total)	-	-	-

Parameter	Estimate	SE	t-value	p-value
Intercept	$\hat{\beta}_1$	$\sqrt{\hat{\sigma}^2\{(X'X)^{-1}_{1,1}\}}$	Estimate/SE	p
x_1	$\hat{\beta}_2$	$\sqrt{\hat{\sigma}^2\{(X'X)^{-1}_{2,2}\}}$	Estimate/SE	p
x_2	$\hat{\beta}_3$	$\sqrt{\hat{\sigma}^2\{(X'X)^{-1}_{3,3}\}}$	Estimate/SE	p

Source	DF	Type I SS	MS	F-value	p-value
x1	1	Type I SS	T1SS/df = T1SS	MS/MSE(model)	p-value
x2	1	Type I SS	T1SS/df = T1SS	MS/MSE(model)	p-value
x3	1	Type I SS	T1SS/df = T1SS	MS/MSE(model)	p-value
		$=\Sigma_{i=1}^{q-1}T1SS_{x_i}$			

Source	Formula	Alternate
CSS(Model)	$\Sigma_{i=1}^{N}\hat{y}_i^2 - N\bar{y}^2 = \Sigma_{i=1}^{N}(\hat{y}_i - \bar{y}^2)$	USS(Model) - SSI
CSS(Total)	$\Sigma_{i=1}^{N}y_i^2 - N\bar{y}^2 = \Sigma_{i=1}^{N}(y_i - \bar{y}^2)$	USS(Total) - SSI & CSS(Model) + SSE
CSS(Error)/SSE	CSS(Total) - CSS(Model)	USS(Total) - USS(model)