Orthogonal Matrix $\Rightarrow$ square mat w/ $A' = A^{-1}$
If sym mat $\Rightarrow \exists$ orth mat V st $A = V\Lambda V'$
$E(AY+B) = AE(Y) + b = A\mu + b$
$Cov(AY+b) = ACov(Y)A' = A\Sigma A'$
$Z^2 = \chi^2(1)$ If $Z_1, \dots, Z_n$ ind:
$W = \sum Z_i^2 \sim \chi^2(n)$
$\Rightarrow \mu = n\sigma^2 = 2n$
$Z/\sqrt{W/n} \sim T(n)$ symmetric about 0
$X_1 \sim \chi^2(n_1) \perp X_2 \sim \chi^2(n_2)$:
$(X_1/n_1)/(X_2/n_2) \sim F(n_1, n_2)$
$T \sim T(v) \quad T^2 \sim F(1, v)$, F is asymmetric, +
$(X_{n\times p})'(X_{n\times p})\hat{\beta} = (X_{n\times p})' y_{n\times 1}$
$(AB)' = B'A' \quad A^{-1}A = AA^{-1} = I$
$(AB)^{-1} = B^{-1}A^{-1} \quad (A')^{-1} = (A^{-1})'$
$|A^{-1}| = \frac{1}{|A|} \quad \frac{1}{|A|}\begin{vmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{vmatrix}$
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
$Cov(Y) = E[(Y-\mu)(Y-\mu)'] = \mathbf{\Sigma}$
$$\begin{vmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \sigma_{n1} & \cdots & \cdots & \sigma_{nn} \end{vmatrix}$$
$\sigma_{ij} = Cov(Y_i, Y_j) = E[(Y_i - \mu_i)(Y_j - \mu_j)']$
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
$X_1, \dots, X_k$ 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 $\Sigma^* = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$
Topic 3, 4 $y = X\beta + \epsilon \quad \hat{\beta} = (X'X)^{-1}X'y$
$\hat{y} = X\hat{\beta} = Hy \quad H = X(X'X)^{-1}X'$
$M = C(X'X)^{-1}C'$

$\hat{\epsilon} = y - \hat{y} = y - Hy = (1 - Hy)$
$\hat{\theta} = C\hat{\beta} \quad \hat{Var}(\hat{\theta}) = \hat{\sigma^2}(C(X'X)^{-1}C')$
$\hat{Var}(\hat{\beta}) = \hat{\sigma^2}(X'X)^{-1}$
$SSH = (\hat{\theta} - \theta_0)'M^{-1}(\hat{\theta} - \theta_0)$
$SSE = (y - \hat{y})'(y - \hat{y}) = y'[I - H]y = \epsilon'\epsilon$
Use MSE for $\hat{\sigma^2}$, use diag $df_{total} = n - 1$
$se = \sqrt{\hat{Var}} \quad MS* = SS*/DF*$
t-value= $\hat{\beta}/se$ p-value= $F(df_{model}, df_{error})$
95% CI $\hat{\theta} \pm 1.96\sqrt{MSE(M)}$
95% PI $\hat{y_h} \pm t(\alpha/2, n-r)se(\hat{y_h})$
$F_{obs} = \frac{(\hat{\theta} - \theta_0)'M^{-1}(\hat{\theta} - \theta_0)/a}{\hat{\sigma^2}} \quad a = rank(C)$
$F_{obs}(1, n-r) = tval^2(n-r)$
One-sided t-Test (only for scalar hypothesis)
$H_A : \theta < \theta_0$ Uses α $H_A : \theta > \theta_0$ uses $(1-\alpha)$
Two-sided t $H_A : \theta \neq \theta_0$ uses $\alpha/2$ and $(1-\alpha/2)$
In all cases reject H_0 if |test stat| > |crit value|
Two-sided F uses α crit value
One-sided F $\theta < \theta_0$ uses 2α and requires appropriate sign of difference
$H_0 : \beta_j = 0$ 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 $t \sim t_{dfE}$
Topic 5: Distributional Results for GLM
X full rank $\hat{\beta} \sim N_p(\beta, \sigma^2(X'X)^{-1}$
$\hat{\theta} \sim N_a(\theta, \sigma^2C(X'X)^{-1}C')$
$\hat{y} = X\hat{\beta} = [X(X'X)^{-1}X']y = Hy$
$E(\hat{y}) = X\beta$
$Cov(\hat{y}) = \sigma^2X(X'X)^{-1}X'$
Resid Var $\Rightarrow \hat{\sigma^2} = \frac{SSE}{n-p} = \frac{\hat{\epsilon}\hat{\epsilon}'}{n-p} = \frac{y'(I-H)y}{n-p}$
Topic 6: MR General Considerations
$R^2_{adj} = 1 - \frac{SSE/(n-r)}{CSS(Total)/(n-1)}$
2 ANOVA tables based on full, reduced:
$F_{obs} = \frac{MSH}{MSE} = \frac{SSH/dfH}{SSE/dfE} = \frac{SSE(reduced) - SSE(full)/dfE(reduced) - dfE(full)}{SSE(full)/dfE(full)}$
$= \frac{CSS(Regression)/(p-1)}{SSE(full)/(n-p)}$
Reject hypothesis if:
$F_{obs} \geq F_F^{-1}(1-\alpha, p-1, n-p) = f_{crit}$
Usual overall regression test assumes model spans an intercept and excludes it from test
$R_c^2 = \frac{CSS(reg)}{CSS(reg) + SSE(full)} = \frac{CSS(reg)}{CSS(total)}$
R_c^2 estimates ρ_c^2, pop ratio of mod to total var
Corrected test for overall regression:
$H_0 : \beta_1 = \cdots = \beta_{p-1} = 0$ holds iff $H_0 : \rho_c^2 = 0$
Topic 7: Testing Hypotheses in MR
All tests compare 2 models:full,reduced (LRT)
Overall test: $F_{obs} = \frac{CSS(\beta_1, \dots, B_{p-1})/(p-1)}{SSE(\beta_0, \dots, \beta_{p-1})/(n-p)}$
added-last test(Type III) : assess the usefulness of one predictor over all others
df always 1 for num b/c testing 1 parameter

F_{obs} = Topic 6 $= \frac{(\hat{\theta} - \theta_0)'M^{-1}(\hat{\theta} - \theta_0)/dfH}{y'(I-H)y/dfE}$
where $C = [0 \dots 0\ 1\ 0 \dots 0]$
Added in order test(Type I): assess the contribution of predictor j over all preceding j-1 predictors without $j+1, \dots$ in model. pval changes based on order of putting things in model unlike Type III
Topic 8: Correlations
$\rho = Corr(X, Y) = \frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$
$R = \frac{\sum(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum(X_i - \bar{X})^2(\sum(Y_i - \bar{y})^2)}}$
Topic 9: GLM assumption Diagnostics
H,L: violations seen in pattern of residuals
I: assessed through logic of sampling scheme
E: finite sample
Gauss: box plot, histogram of resid, test of gauss dist of resid
Discrepancy between T and N vars somewhat inflates prob of reject H_0
Outliers: leverage, Influence $\Rightarrow$ Cook's D
Collinearity:cond ind for kth eigval= $\sqrt{\frac{\lambda_1}{\lambda_k}}$
Condit num: max condit ind
$R_j^2 = R^2(X_j, \{X_1, \dots, X_{j-1}, X_{j+1}, \dots, X_{p-1}\})$
$VIF_j = \frac{1}{1 - R_j^2} = \frac{1}{\text{tolerance}}$

- One-way ANOVA coding schemes:
  Classic ANOVA:
  $y = \mu + \beta_1 I_A + \beta_2 I_B + \beta_3 I_C + \epsilon$
  Cell mean: $y = \beta_1 I_A + \beta_2 I_B + \beta_3 I_C + \epsilon$
  Reference cell: $y = \mu + \beta_1 I_A + \beta_2 I_B + \epsilon$ group C is the reference
- Two-way ANOVA coding schemes:
  Classic ANOVA:
  $y = \mu + \beta_1 I_A + \beta_2 I_B + \beta_3 I_C + \alpha_1 + \alpha_2 + \epsilon$
  Cell means (fit only when they want interaction terms in the model):
  $y = \gamma_{A1} + \gamma_{A2} + \gamma_{A3} + \gamma_{B1} + \gamma_{B2} + \gamma_{B3} + \gamma_{B4} + \epsilon$
  Reference cell: $y = \mu + \beta_1 I_A + \beta_2 I_B + \alpha_1 + \epsilon$

Ref cell: Group 1 is reference μ (grand mean)
Cell mean: all of the β's equal group means
Topic 14: Logistic Regression
$E(y) = p = \frac{\exp(\beta_0 + \beta_1 x)}{1 + \exp(\beta_0 + \beta_1 x)} = \frac{1}{1 + \exp(-\beta_0 - \beta_1 x)}$
$\frac{p}{1-p} = exp(\beta_0 + \beta_1 x)$
$logit(p_i) = \log\left(\frac{p_i}{1-p_i}\right) = \beta_0 + \beta_1 x + \dots$
$= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_{p-1}x_{i,p-1}$ $(x \in 0, 1)$
$OR = \frac{p_1/(1-p_1)}{p_2/(1-p_2)} = \frac{\exp(\beta_0 + \beta_1(1))}{\exp(\beta_0 + \beta_1(0))} =$
$\exp(\beta_0 + \beta_1(1) - \beta_0 - \beta_1(0)) = \exp(\beta_1)$
$\log(OR) = \beta_1$
95% CI OR ($\log(OR) = \hat{\beta}_1$) $\exp(\hat{\beta}_1 \pm 1.96se)$
$y_i \sim Bern(p_i)$ $i = 1, \dots, n$ ys ind of each other
LRT whether k_{th} covariate affects $P(success)$
$H_0 : \beta_k = 0$
Interaction $\Rightarrow$ LRT of full vs reduced
H_0 : interaction not significant
LRT comparing nested models:
$-2\log(L(smaller)) - (-2\log(L(larger)))$
$H_0 \sim \chi_k^2 \quad k = df_{larger} - df_{smaller}$

$H_0 : \beta_j = B_{jk(interaction)} = 0$
$Var = p(1-p)$
Topic 15: Mixed Effect Model
everything is the same as for regression models only difference is covariates
Y=fixed effects + random effects + error
$Y_i + X_i\beta + Z_ib_i + \epsilon_i$
Z_i indicator vars for cluster(family)membership
b_i q-dim vector of random effects
Blocking: group homogeneous exper units together to form block, assign treats at random to exper units within block

Mixed Effect Models 3 types of covariance matrices.
Unstructured: Contains $\frac{k(k+1)}{2}$ unique parameters. LRT tests whether off-diagonals are zero and diagonals are the same. k*(k-1)/2 total parameters, but testing that the diagonal elements are the same. That's why you have (k*(k-1)/2) - 1 degrees of freedom. Null Model Likelihood Ratio Test in SAS.
$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1k} \\ \sigma_{12} & \sigma_{22} & \dots & \sigma_{2k} \\ \dots & \dots & \dots & \dots \\ \sigma_{1k} & \sigma_{2k} & \dots & \sigma_{kk} \end{bmatrix}$$
Compound Symmetry: Model with subject effect (b) and within-subject error (w):
$Y_{ij} = \mathbf{X}_{ij}\beta + b_i + w_{ij}$ with $b_i \sim N(0, \sigma_b^2)$ and $w_{ij} \sim N(0, \sigma_w^2)$. Contains 2 unique parameters. The LRT SAS outputs for the model is testing whether the following matrix is diagonal (if off-diagonal elements are zero) with some σ_w^2. $H_0 : \sigma_b^2 = 0, \sigma_w^2$'s are the same
$$\begin{bmatrix} \sigma_b^2 + \sigma_w^2 & \sigma_b^2 & \dots & \sigma_b^2 \\ \sigma_b^2 & \sigma_b^2 + \sigma_w^2 & \dots & \sigma_b^2 \\ \dots & \dots & \dots & \dots \\ \sigma_b^2 & \sigma_b^2 & \dots & \sigma_b^2 + \sigma_w^2 \end{bmatrix}$$
Autoregressive: Used for time-series/longitudinal data and contains 2 unique parameters.
$$\sigma^2 \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{k-1} \\ \rho & 1 & \rho & \dots & \rho^{k-2} \\ \dots & \dots & \dots & \dots & \dots \\ \rho^{k-1} & \rho^{k-2} & \dots & \rho & 1 \end{bmatrix}$$

- The test to compare two nested mixed effect models is an LRT with test statistic $-2log(red) - -2log(full)$ and follows a χ_{df}^2 with df $= \frac{k(k+1)}{2} - 2$.
- For testing CS vs UN, if you reject the null then you have to use UN.
- AR(1) and CS are nested in UN.

Miscellaneous

- ANOVA has only categorical predictors. ANCOVA has categorical and continuous, and FMiEC has interactions between categorical and continuous, making ANCOVA a special type of FMiEC.
- Cluster design–say 40 mice and 3 measures per mice. The covariance matrix will be block diagonal with each mouse being one block. Everything off-diagonal is zero, but those 40 3x3 matrices are non-zero. So the total number of non-zero elements is 9*40 and number of zero is 120*120 - 360.