Proof 1

$b = a^T Y \quad E[b] = a^T X\beta = \beta \Rightarrow a^T X = I$

$\mathscr{L} = \sigma^2 a^T a - (a^T X - I)\lambda$

$\frac{\partial \mathscr{L}}{\partial a} = 0 \Rightarrow a = \frac{X\lambda}{2\sigma^2} \quad \frac{\partial \mathscr{L}}{\partial \lambda} = 0 \Rightarrow a^T X = I$

$\frac{\lambda^T X^T X}{2\sigma^2} = I \Rightarrow \lambda^T = 2\sigma^2 (X^T X)^{-1}$

$a = X(X^T X)^{-T} \Rightarrow b = (X^T X)^{-1} X^T Y$

Proof 2

$e = (I_n - H)Y \quad MSE = \frac{SSE}{n-p}$

$E[Y^T(I_n - H)Y] = E[(X\beta + \varepsilon)^T (I_n - H)(X\beta + \varepsilon)]$

$\beta^T X^T (I_n - H) X\beta + 2\beta^T X^T (I_n - H) E[\varepsilon] + E[\varepsilon^T (I_n - H)\varepsilon]$

$\cancel{\beta^T X^T (X - X(X^T X)^{-1} X^T X)\beta} + 0 + tr((I_n - H)\sigma^2) + \cancel{E[\varepsilon]^T (I_n - H) E[\varepsilon]}$

$\sigma^2 (tr(I_n) - tr(H))$

$\sigma^2 (tr(I_n) - tr(X(X^T X)^{-1} X^T))$

$\sigma^2 (tr(I_n) - tr(I_p))$

$\sigma^2 (n - p)$

Distribution	**Notes**	**Lower**	**Upper**
$\frac{(n-1)s^2}{\sigma^2} \sim \chi^2_{n-1}$	$s \to MSE \Rightarrow n-1 \to n-p$	$\frac{(n-1)s^2}{\chi^2_{n-1,1-\alpha/2}}$	$\frac{(n-1)s^2}{\chi^2_{n-1,\alpha/2}}$
$\frac{\overline{Y}-\mu}{s/\sqrt{n}} \sim t_{n-1}$	$\overline{Y} \sim N(\mu, \frac{\sigma^2}{n})$	$\overline{Y} - t_{n-1,1-\alpha/2} \frac{s}{\sqrt{n}}$	$\overline{Y} + t_{n-1,1-\alpha/2} \frac{s}{\sqrt{n}}$
$\frac{\overline{Y}_1-\overline{Y}_2-(\mu_1-\mu_2)}{s_p\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n_1+n_2-2}$	$\sigma_1^2 = \sigma_2^2$	$\overline{Y}_1 - \overline{Y}_2 - t_{n_1+n_2-2,1-\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$\overline{Y}_1 - \overline{Y}_2 + t_{n_1+n_2-2,1-\alpha/2} s_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
$\frac{\overline{Y}_1-\overline{Y}_2-(\mu_1-\mu_2)}{\sqrt{\frac{s_1^2}{n_1}+\frac{s_2^2}{n_2}}} \sim t_{df}$	$\sigma_1^2 \neq \sigma_2^2$ df=$\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2 \Big/ \left(\frac{s_1^4}{n_1^2(n_1-1)} + \frac{s_2^4}{n_2^2(n_2-1)}\right)$	$\overline{Y}_1 - \overline{Y}_2 - t_{df,1-\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$	$\overline{Y}_1 - \overline{Y}_2 + t_{df,1-\alpha/2} \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$
$\frac{s_1^2}{s_2^2} \frac{\sigma_2^2}{\sigma_1^2} \sim F_{n_1-1,n_2-1}$	$\sigma_1^2 = \sigma_2^2$ under H_0	$\frac{s_1^2}{s_2^2}/F_{n_1-1,n_2-1,1-\alpha/2},$	$\frac{s_1^2}{s_2^2}/F_{n_1-1,n_2-1,\alpha/2}$
$z' = \frac{1}{2}\log\left(\frac{1+r}{1-r}\right)$ Need CLT, $\hat{\rho} = r$	Pearson $\rho = 0 \Rightarrow \beta_1 = 0$ under H_0, $r = \frac{e^{2z}-1}{e^{2z}+1}$	$z' - Z_{1-\alpha/2}\sqrt{\frac{1}{n-3}}$	$z' + Z_{1-\alpha/2}\sqrt{\frac{1}{n-3}}$
$\frac{\overline{d}_1-\overline{d}_2}{s\sqrt{\frac{1}{n_1}+\frac{1}{n_2}}} \sim t_{n-2}$ Under H_0 need CLT	BF $s^2 = \frac{(\sum_1^{n_1}(d_{i1}-\overline{d}_1+\sum_1^{n_2}(d_{i2}-\overline{d}_2)}{n-2}$ $d_{ij} = \lvert e_{ij} - med(e_j)\rvert$	$\overline{d}_1 - \overline{d}_2 - t_{n-2,1-\alpha/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$	$\overline{d}_1 - \overline{d}_2 + t_{n-2,1-\alpha/2} s \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$
$\frac{SSR^*/2}{(SSE/n)^2} \sim \chi_1^2$ Under H_0	BP $\varepsilon_i \sim N(0, \sigma_i^2) \quad \log\sigma_i^2 = \gamma_0 + \gamma_1 x_i$ $H_0{:}\gamma_1=0 \quad SSR^*$ from $lm(e^2 \sim x)$		
$\frac{b_k-\beta_k}{s(b_k)} \sim t_{n-p}$	$\mathbf{s}^2(\mathbf{b}) = MSE(X^T X)^{-1}$	$b_k - t_{n-p,1-\alpha/2} s(b_k)$	$b_k + t_{n-p,1-\alpha/2} s(b_k)$
$\frac{\widehat{Y}_h-Y_h}{s(\widehat{Y}_h)} \sim t_{n-p}$	$s^2(\widehat{Y}_h) = MSE\, X_h^T (X^T X)^{-1} X_h$	$\widehat{Y}_h - t_{n-p,1-\alpha/2} s(\widehat{Y}_h)$	$\widehat{Y}_h + t_{n-p,1-\alpha/2} s(\widehat{Y}_h)$
$\frac{\widehat{Y}_{h,new}-Y_{h,new}}{s(\widehat{Y}_{h,new})} \sim t_{n-p}$	$s^2(\widehat{Y}_{h,new}) = MSE(1 + X_h^T (X^T X)^{-1} X_h)$	$\widehat{Y}_{h,new} - t_{n-p,1-\alpha/2} s(\widehat{Y}_{h,new})$	$\widehat{Y}_{h,new} + t_{n-p,1-\alpha/2} s(\widehat{Y}_{h,new})$

Source	**SS**	**df**	**MS**	**F-value**	**Pr>F**
Regression	SSR	$p-1$	$MSR = \frac{SSR}{p-1}$	$F^* = \frac{MSR}{MSE}$	p-value
Error	SSE	$n-p$	$MSE = \frac{SSE}{n-p}$		
Total	SSTO	$n-1$			

$H_0 : \beta_1 = \cdots = \beta_{p-1} = 0$

H_a : Not all β_k $'s$ are zero

$F^* \sim F_{p-1,n-p}$

Model	**Hypothesis**	**F-Statistic**
$Y_i = \beta_0 + \beta_1 x_1 + \varepsilon_i$	$H_0 : \beta_1 = 0$	$\frac{SSR(x_1)}{MSE(x_1)}$
$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon_i$	$H_0 : \beta_2 = 0$	$\frac{SSR(x_2\mid x_1)}{MSE(x_1,x_2)}$
$Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon_i$	$H_0 : \beta_3 = 0$	$\frac{SSR(x_3\mid x_1,x_3)}{MSE(x_1,x_2,x_3)}$

$SSR(F) = SSR(x_1, x_2, x_3) = SSR(x_1) + SSR(x_2|x_1) + SSR(x_3|x_1, x_2)$

How much total variation is reduced when adding the other covariates

Can we drop q predictors from the model?

$F^* = \frac{(SSE(R)-SSE(F))/q}{SSE(F)/(n-p)} = \frac{SSR(\{x_i\}_{i=j+1}^k|\{x_i\}_{i=1}^j)/q}{SSE(F)/(n-p)} \sim F_{q,n-p}$

$H_0 : \beta_{i_1} = \cdots = \beta_{i_q} = 0$

H_a : Not all are 0

ANOVA	**df**	**Sun Sq**	**Mean Sq**	**F-Statistic**	**p-value**
x_1	1	$SSR(x_1)$	$SSR(x_1)$	$\frac{SSR(x_1)}{MSE(F)} \sim F_{1,n-2}$	p-value(x_1)
x_2	1	$SSR(x_2\mid x_1)$	$SSR(x_2\mid x_1)$	$\frac{SSR(x_2\mid x_1)}{MSE(F)} \sim F_{1,n-3}$	p-value($x_2\mid x_1$)
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	
x_{p-1}	1	$SSR(F)$	$SSR(F)$	$\frac{SSR(F)}{MSE(F)} \sim F_{1,n-p}$	p-value(F)
Residuals	$n-p$	$SSE(F)$	$MSE(F) = \frac{SSE(F)}{n-p}$		(p-value(F) is same at t p-value(F))

$SSTO = \sum Sum\ Sq$

$(SSE(R) - SSE(F)) =$

$SSR(\{x_i\}_{i=j+1}^k|\{x_i\}_{i=1}^j)$