1) a) $P(\exists j = 1, \dots, m$ such that $H_{0,j}$ is true and $p_j \le \alpha/m)$

$= P\left\{\bigcup_{j=1}^{m_0}\left(p_j \le \frac{\alpha}{m}\right)\right\}$ (suppose there are m_0 true H_0 $\Rightarrow \pi_0 = \frac{m_0}{m}$)

$\le \sum_{j=1}^{m_0} P\left(p_j \le \frac{\alpha}{m}\right)$ (Boole's inequality)

$= \frac{\alpha}{m} \cdot m_0 = \alpha\left(\frac{m_0}{m}\right) = \alpha\pi_0$

1b) $$P\left(\bigcup_{j=1}^{m_0}\left(p_j \le \frac{\alpha}{m}\right)\right) = \sum_{j=1}^{m_0} P\left(p_j \le \frac{\alpha}{m}\right) - \sum_{j<k} P\left[\left(p_j \le \frac{\alpha}{m}\right)\cap\left(p_k \le \frac{\alpha}{m}\right)\right]$$

$$+ \sum_{j<k<l} P\left[\left(p_j \le \frac{\alpha}{m}\right)\cap\left(p_k \le \frac{\alpha}{m}\right)\cap\left(p_l \le \frac{\alpha}{m}\right)\right]$$

$$+ \dots + (-1)^{m_0-1} \sum_{j<\dots<m_0} P\left(\bigcap_{j=1}^{m_0}\left(p_j \le \frac{\alpha}{m}\right)\right)$$

When the p_j's are independent

$$\sum_{j<k} P\left(\left(p_j \le \frac{\alpha}{m}\right)\cap\left(p_k \le \frac{\alpha}{m}\right)\right) = \sum_{j<k} P\left(p_j \le \frac{\alpha}{m}\right) P\left(p_k \le \frac{\alpha}{m}\right) = \binom{m_0}{2}\left(\frac{\alpha}{m}\right)^2$$

$$\sum_{j<k<l} P\left(\left(p_j \le \frac{\alpha}{m}\right)\cap\left(p_k \le \frac{\alpha}{m}\right)\cap\left(p_l \le \frac{\alpha}{m}\right)\right) = \binom{m_0}{3}\left(\frac{\alpha}{m}\right)^3$$

$\dots$

$$\sum_{j<\dots<m_0} P\left(\bigcap_{j=1}^{m_0}\left(p_j \le \frac{\alpha}{m}\right)\right) = \binom{m_0}{m_0}\left(\frac{\alpha}{m}\right)^{m_0} = \left(\frac{\alpha}{m}\right)^{m_0}$$

$$\Rightarrow P\left(\bigcup_{j=1}^{m_0}\left(p_j \le \frac{\alpha}{m}\right)\right) = m_0\frac{\alpha}{m} - \binom{m_0}{2}\left(\frac{\alpha}{m}\right)^2 + \binom{m_0}{3}\left(\frac{\alpha}{m}\right)^3 + \dots + (-1)^{m_0-1}\left(\frac{\alpha}{m}\right)^{m_0}$$

$$= \frac{m_0}{m}\alpha\left(1 - \binom{m_0}{2}\frac{\alpha}{m\,m_0} + \binom{m_0}{3}\frac{\alpha^2}{m_0 m^2} + \dots + \frac{(-1)^{m_0-1}}{m_0}\left(\frac{\alpha}{m}\right)^{m_0-1}\right)$$

$$= \pi_0 \alpha \{1 + o(1)\}$$

where $o(1)$ here is $\sum_{i=2}^{m_0} (-1)^{i-1} \binom{m_0}{i} \frac{1}{m_0} \left(\frac{\alpha}{m}\right)^{i-1}$

This term goes to 0 as α goes to 0.

1c) When the p_j's are dependent

$$P\left(\bigcup_{j=1}^{m_0} (p_j \leq \tfrac{\alpha}{m})\right) = \sum_{j=1}^{m_0} P(p_j \leq \tfrac{\alpha}{m}) + \sum_{i=2}^{m_0} (-1)^{i-1} \sum_{\substack{I \subseteq \{2,\dots,m_0\} \\ |I| = i}} P\left(\bigcap_{i=2}^{m_0} (p_i \leq \tfrac{\alpha}{m})\right)$$

However, unlike independent situation, we <u>don't</u> have

$$P\left(\bigcap_i (p_i \leq \tfrac{\alpha}{m})\right) = \prod_i P(p_i \leq \tfrac{\alpha}{m}) \quad \text{anymore}$$

$\Rightarrow$ We cannot factor out $\left(m_0 \frac{\alpha}{m}\right)$ from all of the terms here

For example $P\left((p_j \leq \tfrac{\alpha}{m}) \cap (p_k \leq \tfrac{\alpha}{m})\right)$

$$= P(p_j \leq \tfrac{\alpha}{m})\, P(p_k \leq \tfrac{\alpha}{m} \mid p_j \leq \tfrac{\alpha}{m})$$

which we are not sure what the form will be since the p-values p_j and p_k are correlated

Therefore, the result from part (b) doesn't hold for dependent p-values

2) a) Let v_i be the eigenvalues of V $(i = 1, \dots, n)$

and a_i be the eigenvalues of A

$$\det(V) = \prod_{i=1}^{n} v_i \quad , \quad \det(A) = \prod_{i=1}^{n} a_i \quad , \quad \det(AV^{-1}) = \prod_{i=1}^{n} \frac{a_i}{v_i}$$

$$\log(\det(V + \epsilon A)) - \log\{\det(V)\} = \log(\det(V+\epsilon A)\det(V^{-1}))$$

$$= \log(\det[(V + \epsilon A)V^{-1}]) = \log(\det(I + \epsilon AV^{-1}))$$

Look at $\det(I + \epsilon AV^{-1}) = \prod_{i=1}^{n} (1 + \epsilon \frac{a_i}{v_i})$

$$= 1 + \epsilon \sum_{i=1}^{n} \frac{a_i}{v_i} + O(\epsilon) = 1 + \epsilon \operatorname{Tr}(AV^{-1}) + O(\epsilon)$$

$$\Rightarrow \log(\det(I + \epsilon AV^{-1})) = \log(1 + \epsilon \operatorname{Tr}(AV^{-1}) + O(\epsilon))$$

$$\approx \epsilon \operatorname{Tr}(AV^{-1}) + O(\epsilon)$$ (Taylor's series' expansion assuming ϵ is small)

$$u^T(V+\epsilon A)^{-1}u - u^T V^{-1} u$$

$$= u^T\left[(V+\epsilon A)^{-1} - V^{-1}\right] u$$

$$= u^T\left[(V+\epsilon A)^{-1}(V+\epsilon A - V)(-V^{-1})\right] u$$

$$= u^T\left[(V+\epsilon A)^{-1}(-\epsilon A V^{-1})\right] u$$

$$= u^T\left[\left(V^{-1} - (V + V\epsilon A^{-1} V)^{-1}\right)(-\epsilon A V^{-1})\right] u$$

$$= -\epsilon u^T V^{-1} A V^{-1} u + \underbrace{\epsilon u^T (V + V\epsilon A^{-1} V)^{-1} A V^{-1} u}_{}$$

$$= -\epsilon u^T V^{-1} A V^{-1} u + o(\epsilon)$$

2b) $$\Sigma(\theta+\delta) - \Sigma(\theta) = \sum_{j=1}^{p} \delta_j M_j(\theta) + o(\|\delta\|_2)$$

$$l(\theta) = -\frac{1}{2}\log\left[\det\{\Sigma(\theta)\}\right] - \frac{1}{2} Y^T\{\Sigma(\theta)\}^{-1} Y$$

$$\nabla_\theta l(\theta) = -\frac{1}{2}\mathrm{Tr}\left(\Sigma(\theta)^{-1}\frac{\partial \Sigma(\theta)}{\partial\theta}\right) + \frac{1}{2} Y^T \Sigma(\theta)^{-1}\frac{\partial\Sigma(\theta)}{\partial\theta}\Sigma(\theta)^{-1} Y$$

$$\Rightarrow (\nabla_\theta l(\theta))_j = -\frac{1}{2}\mathrm{Tr}\left(\Sigma(\theta)^{-1} M_j(\theta)\right) + \frac{1}{2} Y^T (\Sigma(\theta))^{-1} M_j(\theta)\Sigma(\theta)^{-1} Y$$

$$\Rightarrow E(\nabla_\theta l(\theta))_j = -\frac{1}{2}\mathrm{Tr}\left(\Sigma(\theta)^{-1} M_j(\theta)\right) + \frac{1}{2} E(\cancel{Y})^{T\,0}\left(\Sigma(\theta)^{-1} M_j(\theta)\Sigma(\theta)^{-1}\right) \cdot E(\cancel{Y})^{0}$$

$$+ \frac{1}{2}\mathrm{Tr}\left(\Sigma(\theta)^{-1} M_j(\theta)\right)$$

$$= 0$$

$$\Rightarrow E(\nabla_\theta l(\theta)) = 0$$

This result holds using the assumptions $E(Y) = 0$, $Var(Y) = \Sigma(\theta)$. We don't need Y to be normally distributed for this result.

Scanned with CamScanner

HW7

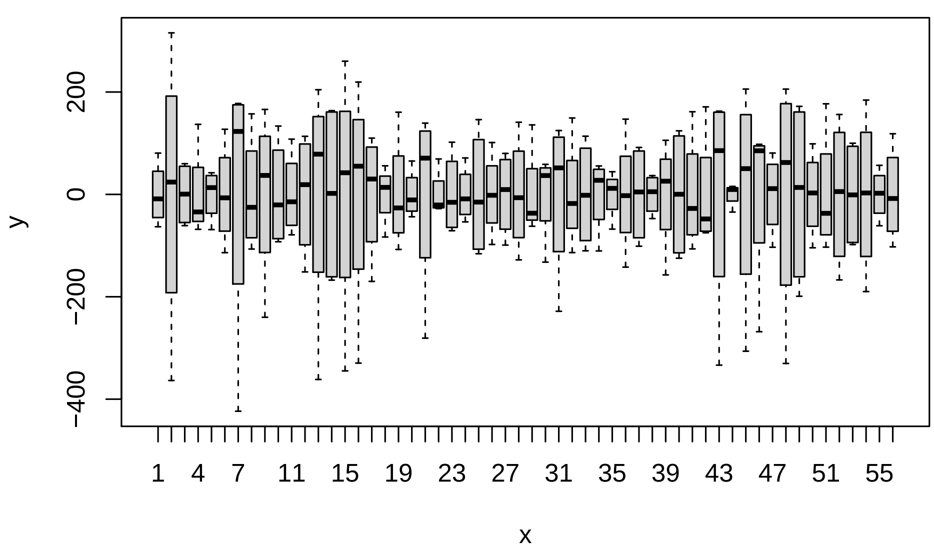
Giang Vu

3/22/2021

3.

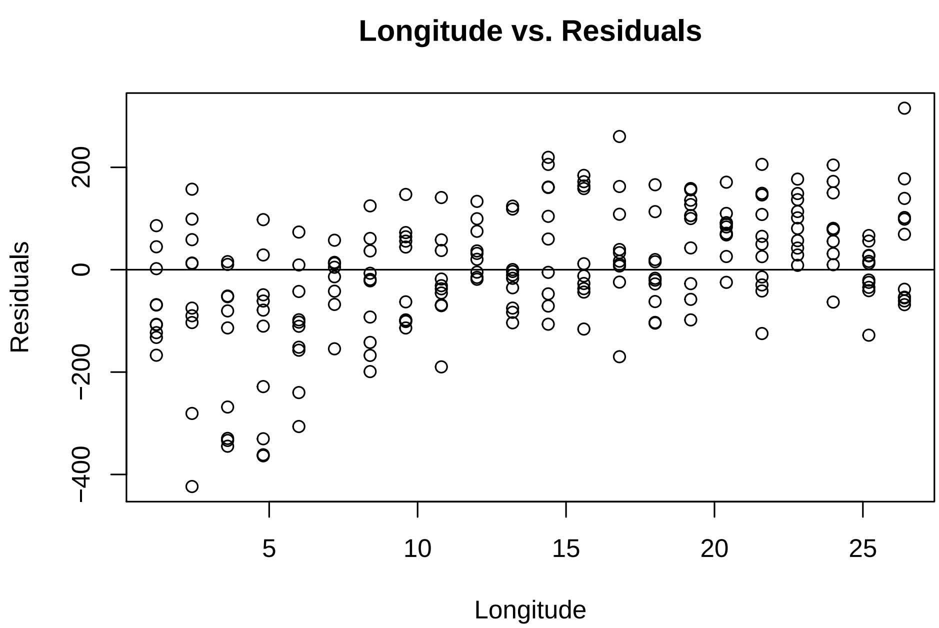
(a) I fitted the standard model for an RCBD treating blocks and varieties as fixed effects, and the coefficient estimate for Variety isn't statistically significant.
That means that there are no significant differences among yield of different varieties.
Included below is also the residual plot.
The p-value for the F test for the hypothesis of no variety effects is 0.7119, which is quite large, so we can't reject this hypothesis.

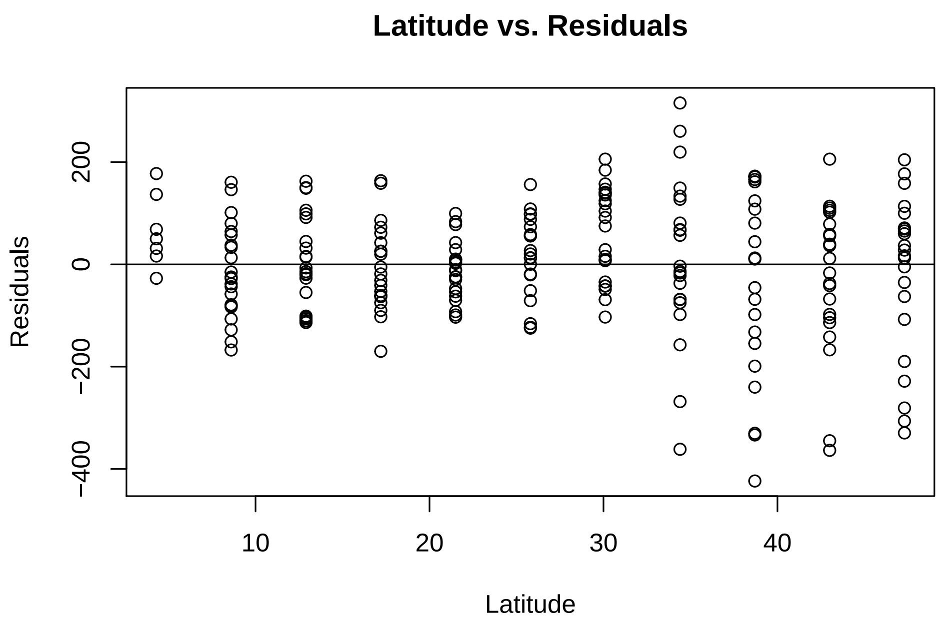
```
## Analysis of Variance Table
##
## Response: Yield
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## Variety    55  954995   17364  0.8755    0.7119
## block       3  723630  241210 12.1621 3.127e-07 ***
## Residuals 165 3272436   19833
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



(b) The estimation of the difference between two means is
$\hat{D} = 8.813$
The standard error of this estimate is
$s^2\{\hat{D}\} = 39.275$
With $r = 56$ and $n_T - r = 224 - 56 = 168$, we obtained thee percentile for the studentized range distribution
$q(.95; 56, 168) = 5.835$
Therefore $T = \frac{1}{\sqrt{2}} q(.95; 56, 168) = 4.126$
The 95% CI for the mean yield of varieties 1-20 minus the mean yield of varieties 21-56 is $\hat{D} \pm Ts\{\hat{D}\}$, which is (-153.2498, 170.8748).

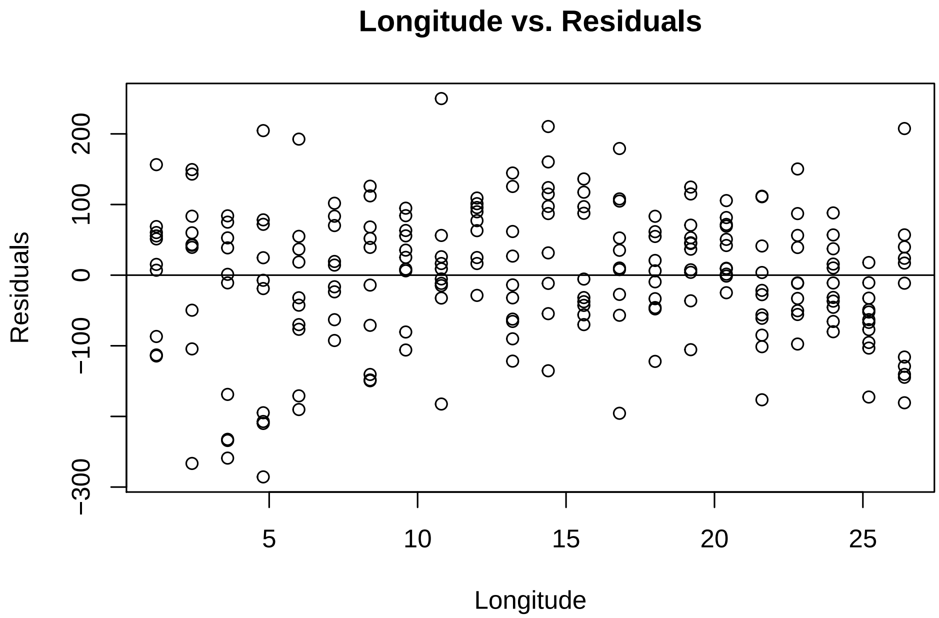
(c) Below is the residual plots versus longitude and latitude. I noticed there are some certain patterns. Lower longitudes tend to have more negative residuals, so therefore have underestimated yields.
On the other hand, higher latitudes have more negative residuals so varieties with higher latitude have more underestimated yields.
This could mean that the model we are assuming might not be good enough, and we might need to include Longitude and Latitude in our model.

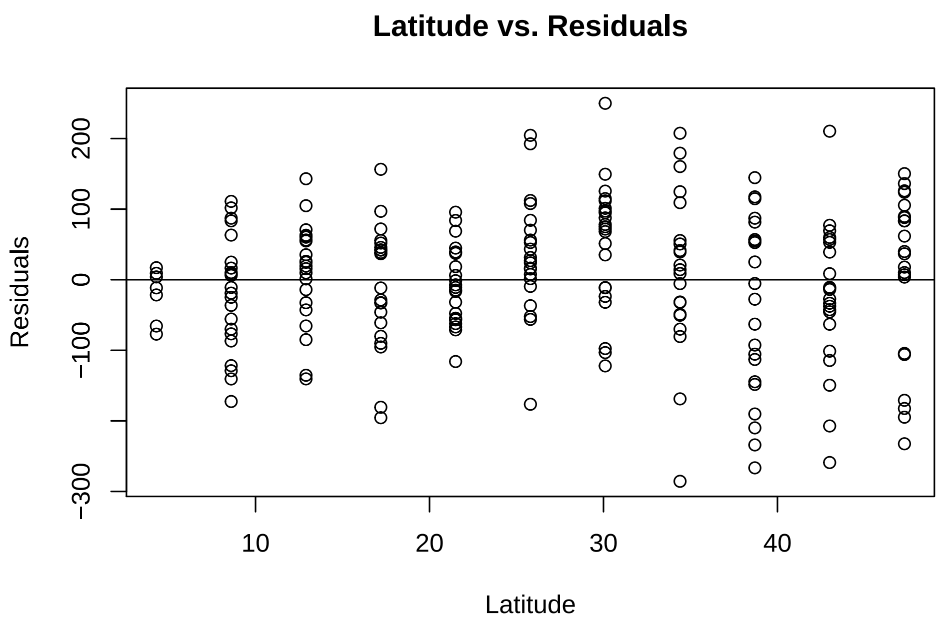




(d) When we added a linear function of the coordinates in our model, the p value for variety effects is smaller but is still not statistically significant at level 0.05.
I also noticed the standard errors of each variety's estimated effect are reduced a lot compared to the original model.
When we plot the residuals as a function of the geographic coordinates of the plots, the patterns identified in part (c) seem to disappear, as the residuals are more evenly scattered across the plots, regardless of the size of longitude and latitude.

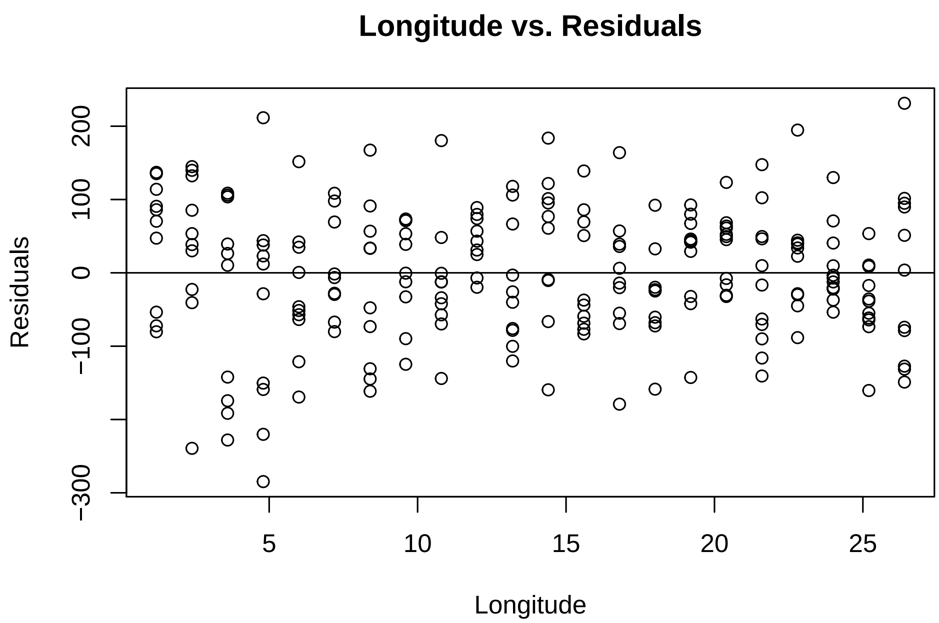
```
## Analysis of Variance Table
##
## Response: Yield
##            Df  Sum Sq Mean Sq F value    Pr(>F)
## Variety    55  954995   17364  1.3883   0.05954 .
## block       3  723630  241210 19.2853 9.436e-11 ***
## Latitude    1  305482  305482 24.4240 1.904e-06 ***
## Longitude   1  928239  928239 74.2148 5.769e-15 ***
## Residuals 163 2038716   12507
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

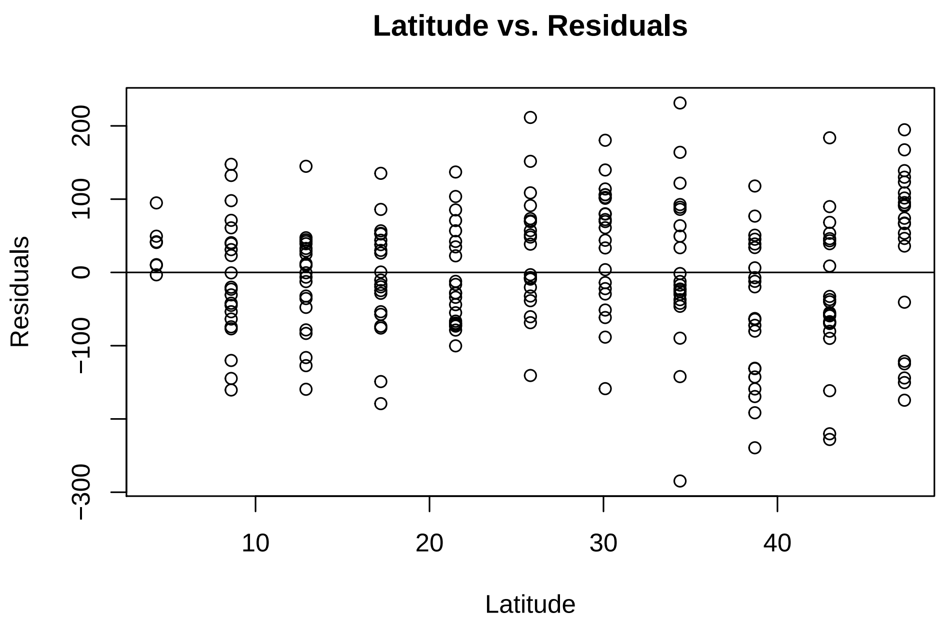




(e) When we added a second order polynomial of the coordinates in our model, the p value for variety effects is much smaller and now is statistically significant at level 0.05.
The standard errors of each variety's estimated effect are reduced even more compared to original model and model in part (d).
When we plot the residuals as a function of the geographic coordinates of the plots, the patterns identified in part (c) seem to disappear as well, as the residuals are more evenly scattered across the plots, regardless of the size of longitude and latitude, just like residual plots in part (d).

```
## Analysis of Variance Table
##
## Response: Yield
##                  Df  Sum Sq Mean Sq F value    Pr(>F)
## Variety          55  954995   17364  1.5438  0.019576 *
## block             3  723630  241210 21.4463 9.724e-12 ***
## Latitude          1  305482  305482 27.1608 5.668e-07 ***
## Longitude         1  928239  928239 82.5310 3.628e-16 ***
## I(Latitude^2)     1  115815  115815 10.2973  0.001608 **
## I(Longitude^2)    1  112109  112109  9.9678  0.001902 **
## Residuals       161 1810792   11247
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```





(f) After carrying out all the above analyses, I think this study should be redesigned so that the main factor is a new variable that contains information about both variety and geographical coordinates.
With the original model, we saw no variety effects when we fit it, and that problem went away when we included coordinates into our model.
That means the yields of wheat might be explained by both variety and geographical location.

4) a) For the one-way ANOVA model with p treatmeants

$$Y_{ij} = \mu_i + \epsilon_{ij} \qquad i = 1, \dots, p \text{ and } j = 1, \dots, n_i$$

$$\beta = (\mu_1, \dots, \mu_p)^T$$

$$\Rightarrow \hat{\beta} = (\hat{\mu}_1, \dots, \hat{\mu}_p)^T = (\bar{Y}_{1\cdot}, \dots, \bar{Y}_{p\cdot})^T$$

$$\Rightarrow se(c^T\hat{\beta}) = \sqrt{MSE \sum_{i=1}^{p} \frac{c_i^2}{n_i}}$$

$$c^T\hat{\beta} - c^T\beta = \sum_{i=1}^{p} c_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})$$

Using Cauchy-Schwartz, we have

$$\sum_{i=1}^{p} c_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot}) \leq \sqrt{\sum_i c_i^2 \sum_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2}$$

$$= \sqrt{\sum_i \frac{c_i^2}{n_i} \sum_i n_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})^2}$$

$$= \sqrt{\sum_i \frac{c_i^2}{n_i} (SSTR)} \qquad \text{(from last HW)}$$

$$\Rightarrow \frac{\sum_i c_i (\bar{y}_{i\cdot} - \bar{y}_{\cdot\cdot})}{\sqrt{MSE \sum_i \frac{c_i^2}{n_i}}} \leq \frac{\sqrt{\sum \frac{c_i^2}{n_i} (SSTR)}}{\sqrt{MSE \sum \frac{c_i^2}{n_i}}} = \sqrt{\frac{SSTR}{MSE}}$$

$$\Rightarrow \left\{ \frac{c^T\hat{\beta} - c^T\beta}{se(c^T\hat{\beta})} \right\}^2 \leq \frac{SSTR}{MSE} = \frac{(p-1) MSTR}{MSE}$$

When $y \sim N(X\beta, \sigma^2 I_n) \Rightarrow \frac{MSTR}{MSE} \sim F_{p-1, n_T - p}$

$$\Rightarrow \left\{ \frac{c^T\hat{\beta} - c^T\beta}{se(c^T\hat{\beta})} \right\}^2 \leq (p-1) F_{p-1, n_T - p}$$

We also have $Pr(F \leq \underbrace{F_{(1-\alpha), p-1, n_T-p}}_{(1-\alpha) \text{ quantile}}) = 1 - \alpha$

$$\Rightarrow Pr\left(\left\{ \frac{c^T\hat{\beta} - c^T\beta}{se(c^T\hat{\beta})} \right\}^2 \leq (p-1) F^{(1-\alpha)}_{p-1, n_T-p} \quad \forall c \in S \right) \leq Pr\left((p-1) F_{p-1, n_T-p} \leq (p-1) F^{(1-\alpha)}_{p-1, n_T-p} \right)$$

$$= 1 - \alpha$$

$$\Rightarrow \quad Pr\left[\left\{\frac{c^T\beta - c^T\beta}{se(c^T\hat{\beta})}\right\}^2 \leq q_{1-\alpha} \quad \forall c \in S\right] \lesssim 1-\alpha$$

4b) The $1-\alpha$ CI for every point in $\{c^T\beta : c \in S\}$

is just $\quad c^T\hat{\beta} \pm \sqrt{q_{1-\alpha}}\, se(c^T\hat{\beta})$