

Problem 1:

Part (a):

The residual plot shows there are no bad patterns in the residual plot.

The dotplot suggests there is variation among sites is larger and variation within sites (from plot to plot) appears smaller. In another word, parcels on different sites will be relatively less similar, while parcels on the same site will be relatively similar.

R code for part(a)

```
corn<-read.table("corn.txt", header=T)
attach(corn)
library(lattice)
par(mfrow=c(1,2))
xyplot(ears~site, pch=16, col="red", main="dotplot")
plot(fitted(ears.lmer), residuals(ears.lmer), main="residual plot")
abline(h=0)
```

Part (b):

Model: $y_{ij} = \mu + \alpha_i(\text{site}, \text{random}) + \beta_{ij}(\text{parcel}, \text{random})$

$i = 1, \dots, 8; j = 1, \dots, 4$

The grand mean is 41.219

Standard deviation between sites is $\sigma_L = 6.76$; standard deviation in the same site is $\sigma_W = 3.98$.

Here: $\frac{\sigma_L^2}{\sigma_L^2 + \sigma_W^2} = \frac{45.695}{45.695 + 15.839} = 0.74$. Roughly 74% of the thickness variation is due to differences between sites. This proves our discussion in part (a).

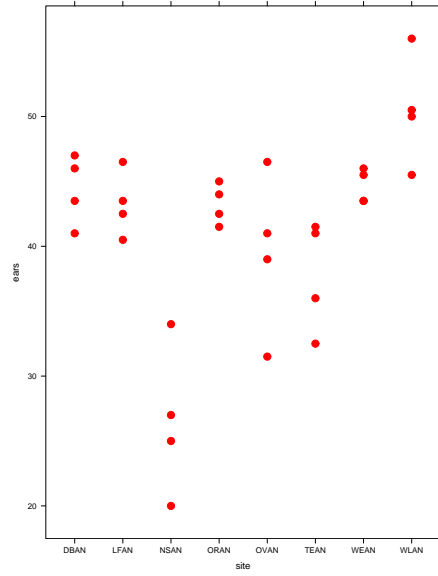


Figure 1: dotplot of the number of ears by site for Problem1(a)

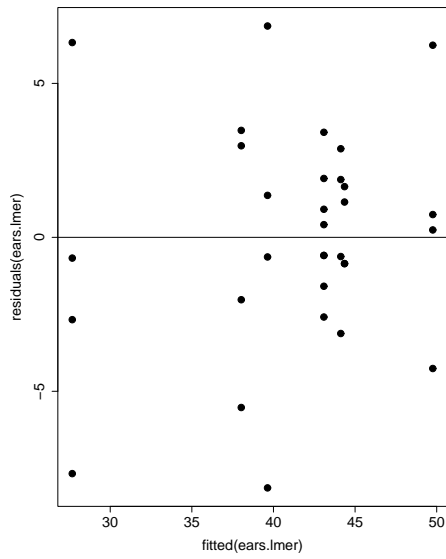


Figure 2: residual plot for Problem1(a)

The sample mean and fitted value for each site:

	mean	fitted
DBAN	44.375	44.12331
LFAN	43.250	43.08802
NSAN	26.500	27.67372
ORAN	43.250	43.08802
OVAN	39.500	39.63706
TEAN	37.750	38.02661
WEAN	44.625	44.35338
WLAN	50.500	49.75989

Rcode and output for part(b):

```
library(lme4)
ears.lmer = lmer(ears~1+(1|site))
summary(ears.lmer)
Linear mixed-effects model fit by REML
Formula: ears ~ 1 + (1 | site)
    AIC   BIC logLik MLdeviance REMLdeviance
198.8 201.7 -97.39    198.4      194.8
Random effects:
Groups   Name          Variance Std.Dev.
site    (Intercept) 45.695   6.7598
Residual                15.839   3.9798
number of obs: 32, groups: site, 8
Fixed effects:
              Estimate Std. Error t value
(Intercept)  41.219      2.491    16.55
```

```

>means=sapply(split(ears, site), mean)
>siteFit=sapply(split(fitted(ears.lmer), site), mean)
>data.frame(mean=means, fitted=siteFit)

```

Part (c):

We assume that corns at parcel V were grown under the same conditions as other corns at site WLAN. From part(b), the number of ears of corn in this parcel is 49.76 and standard error is $\hat{\sigma}_W = 3.98$

Part (d):

The variance of the difference between the two parcels is $\sqrt{2 * \hat{\sigma}_W^2} = \sqrt{2 * 15.839} = 5.63$

Part (e):

$$\sqrt{2 * \hat{\sigma}_W^2} = \sqrt{2 * 15.839} = 5.63$$

Part (f):

Estimate the mean number of ears of corn on average in the six parcels in a new site (with a standard deviation)

The standard deviation is: $\sqrt{\sigma_L^2 + \frac{\sigma_W^2}{6}} = \sqrt{45.695 + \frac{15.839}{6}} = 6.95$

Except the estimated standard deviation, we don't have any other informations about the new site. It is reasonable to use the grand mean, 41.219, to estimate the mean number of ears of corn on average in the six parcels in a new site. The 95% CI is : (27.597,54.841)