

RANDOMIZED BLOCK DESIGNS AND THE ANOVA F-TEST

Consider a **randomized block design** (RBD) with k treatments and b blocks. Assume that each block has k experimental units, and that one unit is assigned to each treatment. Let x_{ij} be the measured response for the experimental unit from block j in treatment i and

- sample mean for **treatment** i

$$\bar{x}_i = \frac{1}{b} \sum_{j=1}^b x_{ij} \quad i = 1, \dots, k$$

- sample mean for block j

$$\bar{x}_j^{(b)} = \frac{1}{k} \sum_{i=1}^k x_{ij} \quad j = 1, \dots, b$$

- overall sample mean

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^b x_{ij}$$

- Sum of Squares for Treatments (SST)

$$SST = \sum_{i=1}^k b(\bar{x}_i - \bar{x})^2$$

- Sum of Squares for Blocks (SSB)

$$SSB = \sum_{j=1}^b k(\bar{x}_j^{(b)} - \bar{x})^2$$

- Overall Sum of Squares (SS)

$$SS = \sum_{i=1}^k \sum_{j=1}^b (x_{ij} - \bar{x})^2$$

The following decomposition holds

$$SS = SST + SSB + SSE \quad \therefore \quad SSE = SS - SST - SSB$$

For testing

$$H_0 : \mu_1 = \dots = \mu_k$$

$$H_a : \text{At least two treatment means different}$$

in an RBD, the test statistic is

$$F = \frac{MST}{MSE}$$

where

$$MST = \frac{SST}{k-1} \quad MSE = \frac{SSE}{n-b-k+1}$$

If H_0 is **true**, then $F \sim F_{k-1, n-b-k+1}$, and the rejection region for the test with significance level α is

$$F > F_{\alpha, k-1, n-b-k+1}$$

where F_{α, ν_1, ν_2} is the upper α point of the F-distribution with ν_1 and ν_2 degrees of freedom.

Example: Measurements were made on the amount of sulphur (in parts per million) in soil samples using four different solvents. The soil samples were collected from five different geographical locations in Florida, USA, and represented different soil types.

The **response variable** sulphur level. The single **factor** is the *solvent* and there are $k = 4$ factor levels:

1. Calcium Chloride (CaCl_2)
2. Ammonium Acetate (NH_4OAc)
3. Mono-Calcium Phosphate ($\text{Ca}(\text{H}_2\text{P O}_4)_3$)
4. Water (H_2O)

The *soil* types determine the $b = 5$ **blocks**

1. Troup, Jackson Co. (*Paleudults* soil)
2. Lakeland, Walton Co. (*Quartzipsamments* soil)
3. Leon, Duval Co. (*Haplaquads* soil)
4. Chipley, Jackson Co. (*Quartzipsamments* soil)
5. Norfolk, Alachua Co. (*Paleudults* soil)

The data observed in the study were as follows:

Treatment	Block				
	Troup	Lakeland	Leon	Chipley	Norfolk
CaCl_2	5.07	3.31	2.54	2.34	4.71
NH_4OAc	4.43	2.74	2.09	2.07	5.29
$\text{Ca}(\text{H}_2\text{P O}_4)_3$	7.09	2.32	1.09	4.38	5.70
H_2O	4.48	2.35	2.70	3.85	4.98

```
> print(summary(aov(sulphur~solvent+soil,data=Soil)), digits = 5)

+               Df Sum Sq Mean Sq F value    Pr(>F)
+ solvent         3  1.621  0.5404  0.6726 0.5851298
+ soil           4 33.965  8.4912 10.5683 0.0006629 ***
+ Residuals     12  9.642  0.8035
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The this is the standard ANOVA table for the randomized block design. As expected, there is a significant difference between **blocks** ($F_{obs} = 10.568$, p-value=0.001), but **no significant difference** between **treatments** ($F_{obs} = 0.673$, p-value=0.585).

THE NEED FOR BLOCKING IN AN RBD ANALYSIS

Consider the following response data: five measurements collected in three treatment groups:

	Block				
—	1	2	3	4	5
Group 1	-20.88	-4.76	-0.46	10.78	-10.47
Group 2	-15.75	0.11	5.64	16.98	-6.03
Group 3	-8.62	6.20	10.42	20.05	-1.29

Raw data broken down by block

```
> Y<-c(-20.88,-4.76,-0.46,10.78,-10.47,
+      -15.75,0.11,5.64,16.98,-6.03,
+      -8.62,6.20,10.42,20.05,-1.29)
>
> Group<-gl(3,5,labels=c("Group1","Group2","Group3"))
> Block<-gl(5,1,15,labels=c(1:5))
> par(mar=c(4,4,1,2))
> gplots::plotmeans(Y~Group,ylim=range(-40,40))
> stripchart(Y~Group,vertical=T,pch=4,cex=0.75,lwd=2,add=T)
```

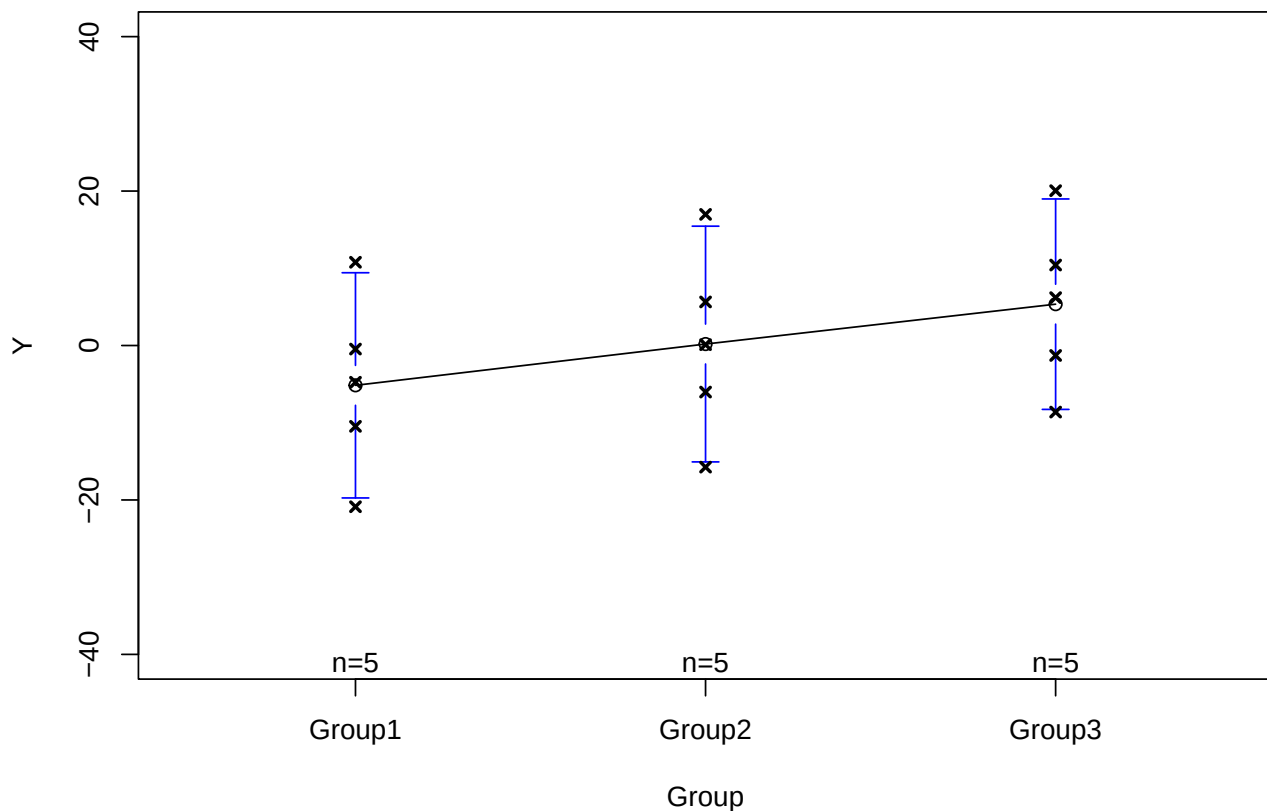


Figure 11: Raw data with block ignored

```
> print(summary(aov(Y~Group)), digits = 5)

+           Df  Sum Sq Mean Sq F value Pr(>F)
+ Group       2   276.18   138.09   1.0118 0.3926
+ Residuals   12  1637.80   136.48
```

It seems from this analysis that there is no difference between groups; however, this is not the case if the block structure is taken into account.

```
> blocklabel<-paste(rep("Block", 5), c(1:5))
> par(mar=c(4, 4, 1, 2))
> gplots::plotmeans(Y~Group, ylim=range(-40, 40))
> points(rep(0.9, 5), Y[Group=="Group1"], pch=c(1:5))
> points(rep(1.9, 5), Y[Group=="Group2"], pch=c(1:5))
> points(rep(2.9, 5), Y[Group=="Group3"], pch=c(1:5))
> legend(0.5, 40, blocklabel, pch=(1:5))
```

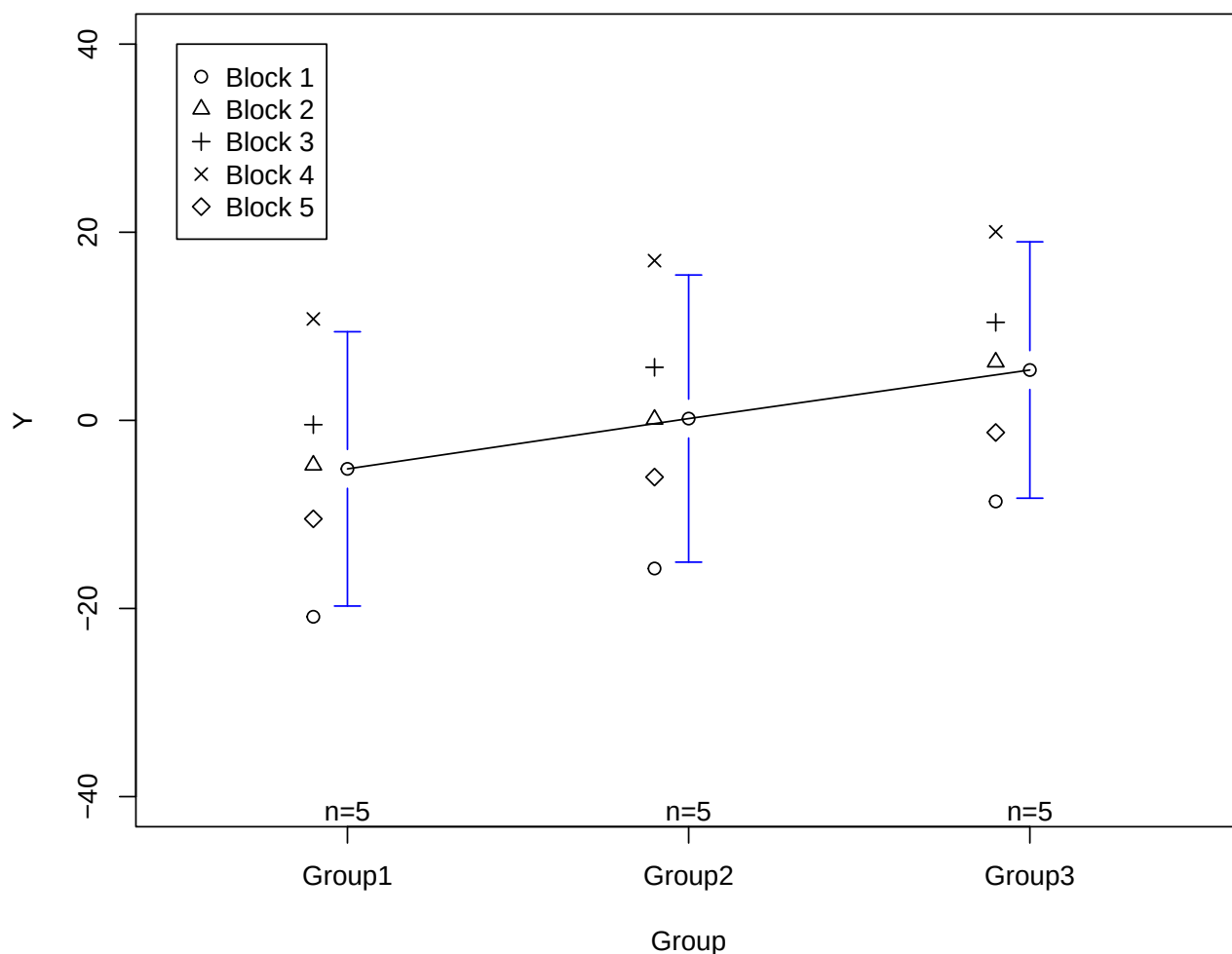


Figure 12: Raw data broken down by block

```
> print(summary(aov(Y~Group+Block)), digits = 5)

+           Df   Sum Sq Mean Sq F value    Pr(>F)
+ Group         2    276.18   138.09   178.81 2.292e-07 ***
+ Block         4   1631.62   407.91   528.18 1.009e-09 ***
+ Residuals     8     6.18     0.77
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

In fact, there is hidden structure in the data. If this structure is taken into account, evidence that the treatment means are significantly different is uncovered. The reason that the CRD and one-way ANOVA do not discover this is that they assume that the variability can be decomposed as

$$SS = SST + SSE$$

whereas in fact

$$SS = SST + SSB + SSE$$

that is, the CRD assumes that the random variability that is observed is MUCH LARGER than it actually is. Once the variation due to BLOCKS is taken into account, the ANOVA-F test result for TREATMENTS becomes significant.

RANDOMIZED COMPLETE BLOCK DESIGNS WITH BALANCED REPLICATION

Consider a **randomized block design** (RBD) with k treatments and b blocks, and r **replications**, giving $n = rbk$ observations in total. Let x_{ijt} be the t th replicated observation in the (i, j) th treatment/block combination.

- sample mean for **treatment** i

$$\bar{x}_i = \frac{1}{br} \sum_{j=1}^b \sum_{t=1}^r x_{ijt} \quad i = 1, \dots, k$$

- sample mean for **block** j

$$\bar{x}_j^{(B)} = \frac{1}{kr} \sum_{i=1}^k \sum_{t=1}^r x_{ijt} \quad j = 1, \dots, b$$

- sample mean for replicates in (i, j) th **treatment/block** combination

$$\bar{x}_{ij} = \frac{1}{r} \sum_{t=1}^r x_{ijt} \quad i = 1, \dots, k, j = 1, \dots, b$$

- overall sample mean

$$\bar{x} = \frac{1}{n} \sum_{i=1}^k \sum_{j=1}^b \sum_{t=1}^r x_{ijt}$$

- Sum of Squares for Treatments (SST)

$$SST = \sum_{i=1}^k br(\bar{x}_i - \bar{x})^2$$

- Sum of Squares for Blocks (SSB)

$$SSB = \sum_{j=1}^b kr(\bar{x}_j^{(B)} - \bar{x})^2$$

- Sum of Squares for Interaction (SSI)

$$SSI = \sum_{i=1}^k \sum_{j=1}^b r(\bar{x}_{ij} - \bar{x}_i - \bar{x}_j^{(B)} + \bar{x})^2$$

- Overall Sum of Squares (SS)

$$SS = \sum_{i=1}^k \sum_{j=1}^b \sum_{t=1}^r (x_{ijt} - \bar{x})^2$$

The following decomposition holds

$$SS = SST + SSB + SSI + SSE \quad \therefore \quad SSE = SS - SST - SSB - SSI$$

Define

$$MST = \frac{SST}{k-1} \quad MSB = \frac{SSB}{b-1} \quad MSI = \frac{SSI}{(k-1)(b-1)}$$

and

$$MSE = \frac{SSE}{n-bk}$$

- For testing for a **TREATMENT** effect, use

$$F = \frac{MST}{MSE}$$

Under the assumption of **NO TREATMENT EFFECT**, then

$$F \sim F_{k-1, n-bk}$$

which defines the rejection region and p-value in the usual way.

- For testing for a **BLOCK** effect, use

$$F = \frac{MSB}{MSE}$$

Under the assumption of **NO BLOCK EFFECT**, then

$$F \sim F_{b-1, n-bk}$$

- For testing for an **INTERACTION**, use

$$F = \frac{MSI}{MSE}$$

Under the assumption of **NO INTERACTION**, then

$$F \sim F_{(k-1)(b-1), n-bk}$$

Example: Measurements were made on the lifetimes of batteries (in hours) for three battery types constructed from different materials, to investigate the effect of operating temperature on lifetime. It was believed before the experiment that the battery types were likely to behave differently in the experiment.

The **response variable** is lifetime. The single **factor** is the *temperature* and there are $k = 3$ factor levels:

1. 15 Celsius
2. 70 Celsius
3. 125 Celsius

Treatment	Block		
	Lead	Acetate	Nickel Cadmium
15	130,155,74,180	150,188,159,126	138,119,168,160
70	34,40,80,75	126,122,106,115	174,120,150,139
120	20,70,82,58	25,70,58,45	96,104,82,60

The *material* types determine the $b = 3$ **blocks**

1. Lead
2. Acetate
3. Nickel Cadmium

$r = 4$ replicate measurements were made, so that $n = 3 \times 3 \times 4 = 36$ data were obtained in total.

The data observed in the study were as follows:

```
> print(summary(aov(life~temp*material,data=Batteries)), digits = 5)

+               Df Sum Sq Mean Sq F value    Pr(>F)
+ temp              2   39083  19541.6  29.3438 1.694e-07 ***
+ material           2   10633   5316.6   7.9834 0.001888 **
+ temp:material      4    9438   2359.4   3.5429 0.018973 *
+ Residuals         27   17981    666.0
+ ---
+ Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> car::leveneTest(life~temp*material,data=Batteries,center='mean')

+ Levene's Test for Homogeneity of Variance (center = "mean")
+      Df F value Pr(>F)
+ group 8    1.059  0.42
+      27

> car::leveneTest(life~temp*material,data=Batteries,center='median')

+ Levene's Test for Homogeneity of Variance (center = "median")
+      Df F value Pr(>F)
+ group 8    0.936  0.504
+      27
```

There is a

- **significant difference between blocks** (material, $F_{obs} = 7.983$, p-value 0.00189),
- **significant difference between treatments** (temp, $F_{obs} = 29.344$, p-value $1.69e - 07$),
- **significant interaction** (temp:material, $F_{obs} = 3.543$, p-value 0.01897).

Levene's test reveals that there is no evidence to suspect that the population variances are different:

```

> par(mar=c(4,4,1,2))
> gplots::plotmeans(life~temp, subset=(material=='Lead'),
+   n.label=FALSE, col='blue', barcol='blue', data=Batteries,
+   ylim=range(0,250))
> suppressWarnings(gplots::plotmeans(life~temp,
+   subset=(material=='Acetate'), n.label=FALSE,
+   col='orange', barcol='orange', data=Batteries, add=TRUE))
> suppressWarnings(gplots::plotmeans(life~temp,
+   subset=(material=='Nickel Cadmium'), n.label=FALSE,
+   col='cyan', barcol='cyan', data=Batteries, add=TRUE))
> legend(2.5,250,c('Lead','Acetate','Nickel Cadmium'),
+   col=c('blue','cyan','orange'), lty=1)

```

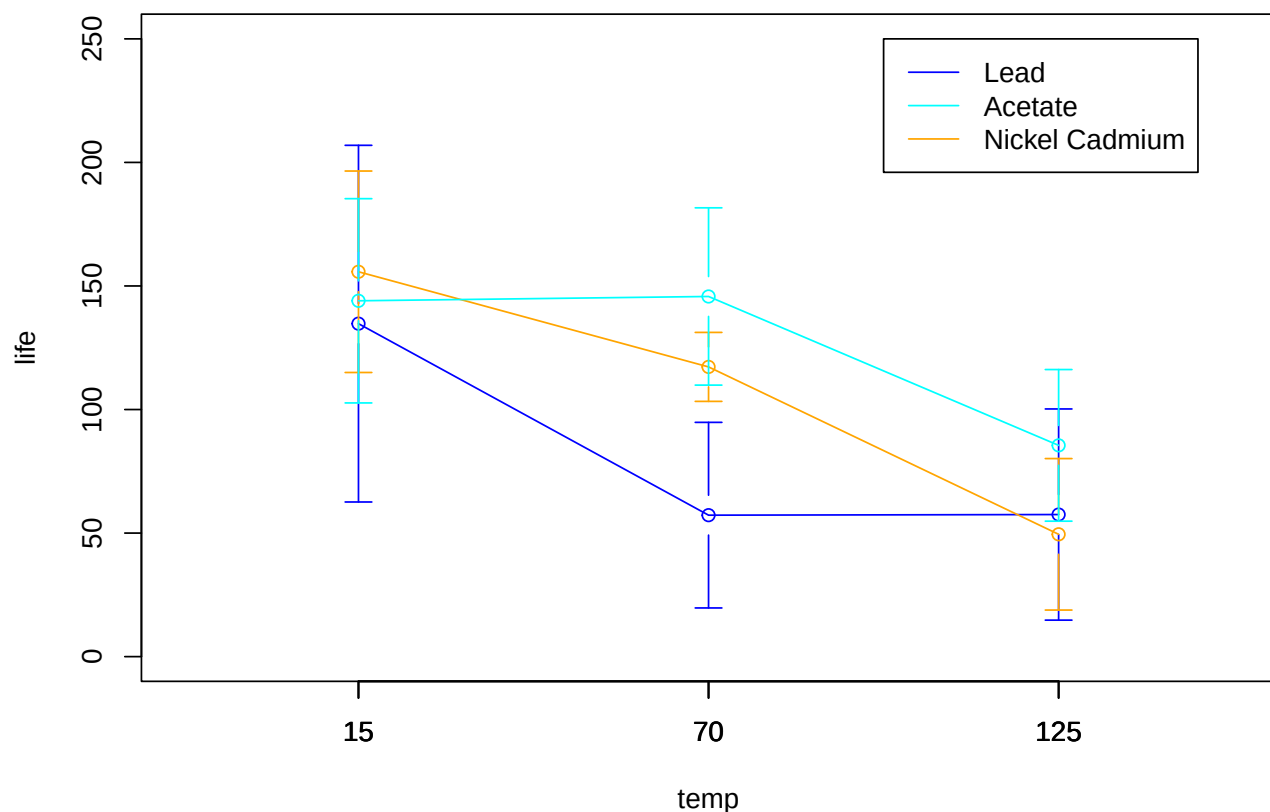


Figure 13: Means Plot by Temperature

```

> par(mar=c(4,4,1,2))
> gplots::plotmeans(life~material, subset=(temp=='15'),
+   n.label=FALSE,
+   col='blue', barcol='blue', data=Batteries,ylim=range(0,250))
> suppressWarnings(
+   gplots::plotmeans(life~material,subset=(temp=='70'),
+   n.label=FALSE, col='orange', barcol='orange', data=Batteries,
+   add=TRUE))
> suppressWarnings(
+   gplots::plotmeans(life~material, subset=(temp=='125'),

```

```

+   n.label=FALSE, col='cyan', barcol='cyan', data=Batteries,
+   add=TRUE) )
> legend(2.5, 250, c('15', '70', '125'),
+   col=c('blue', 'cyan', 'orange'), lty=1)

```

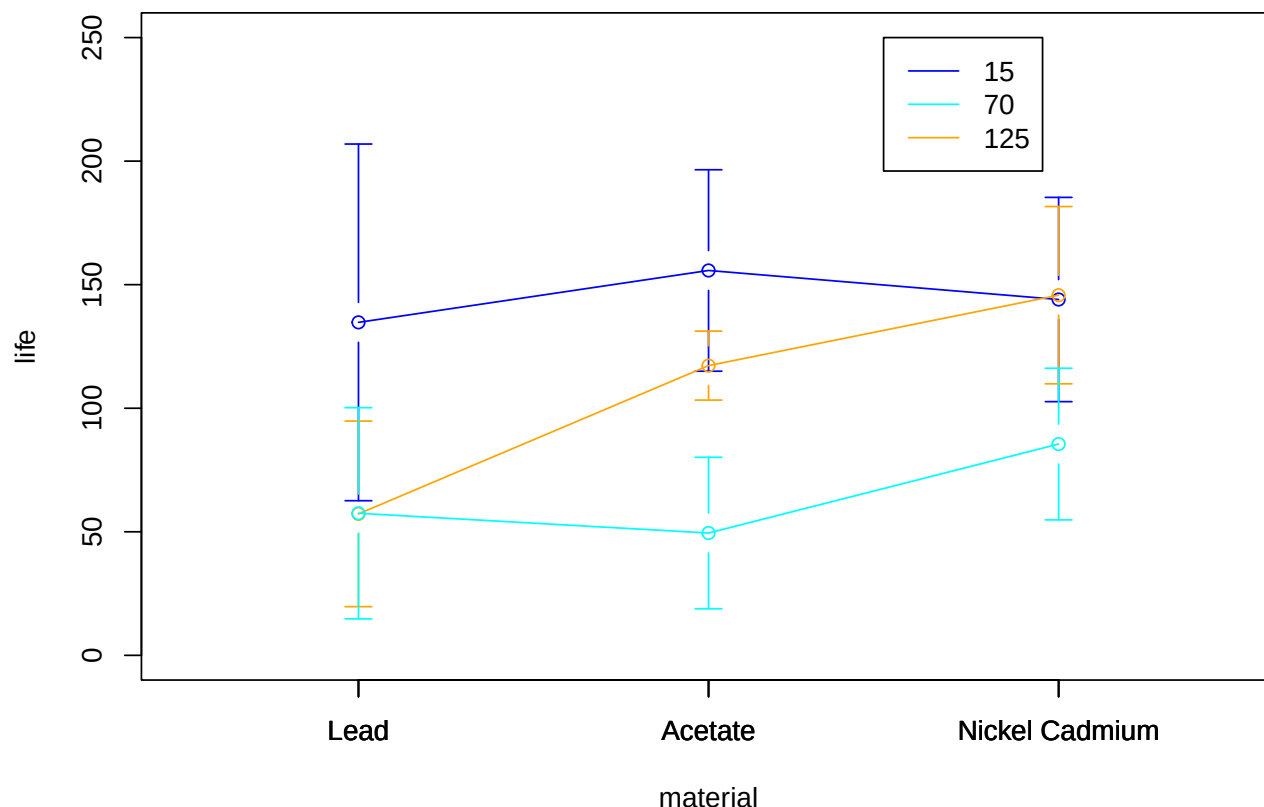


Figure 14: Means Plot by Material

To produce a means plot, we also use the full `ggplot` functionality. First we compute a by-treatment summary using the `summarySE` function from the `Rmisc` library.

```

> par(mar=c(4, 4, 1, 2))
> batsum <- Rmisc::summarySE(Batteries, measurevar="life",
+   groupvars=c("temp", "material"))
> batsum

```

	temp	material	N	life	sd	se	ci	
+	1	15	Lead	4	134.75	45.35324	22.67662	72.1671
+	2	15	Acetate	4	155.75	25.61738	12.80869	40.7630
+	3	15	Nickel Cadmium	4	144.00	25.97435	12.98717	41.3310
+	4	70	Lead	4	57.25	23.59908	11.79954	37.5514
+	5	70	Acetate	4	117.25	8.77021	4.38511	13.9554
+	6	70	Nickel Cadmium	4	145.75	22.54440	11.27220	35.8732
+	7	125	Lead	4	57.50	26.85144	13.42572	42.7266
+	8	125	Acetate	4	49.50	19.26136	9.63068	30.6491
+	9	125	Nickel Cadmium	4	85.50	19.27866	9.63933	30.6766

```

> pd <- position_dodge(0.2)
> ggplot(batsum,
+   aes(x=temp, y=life, colour=material, group=material)) +
+   geom_errorbar(aes(ymin=life-ci, ymax=life+ci),
+     width=0.25, position=pd) +
+   geom_line(position=pd) +
+   geom_point(position=pd, size=3) +
+   theme(legend.position="bottom")

```

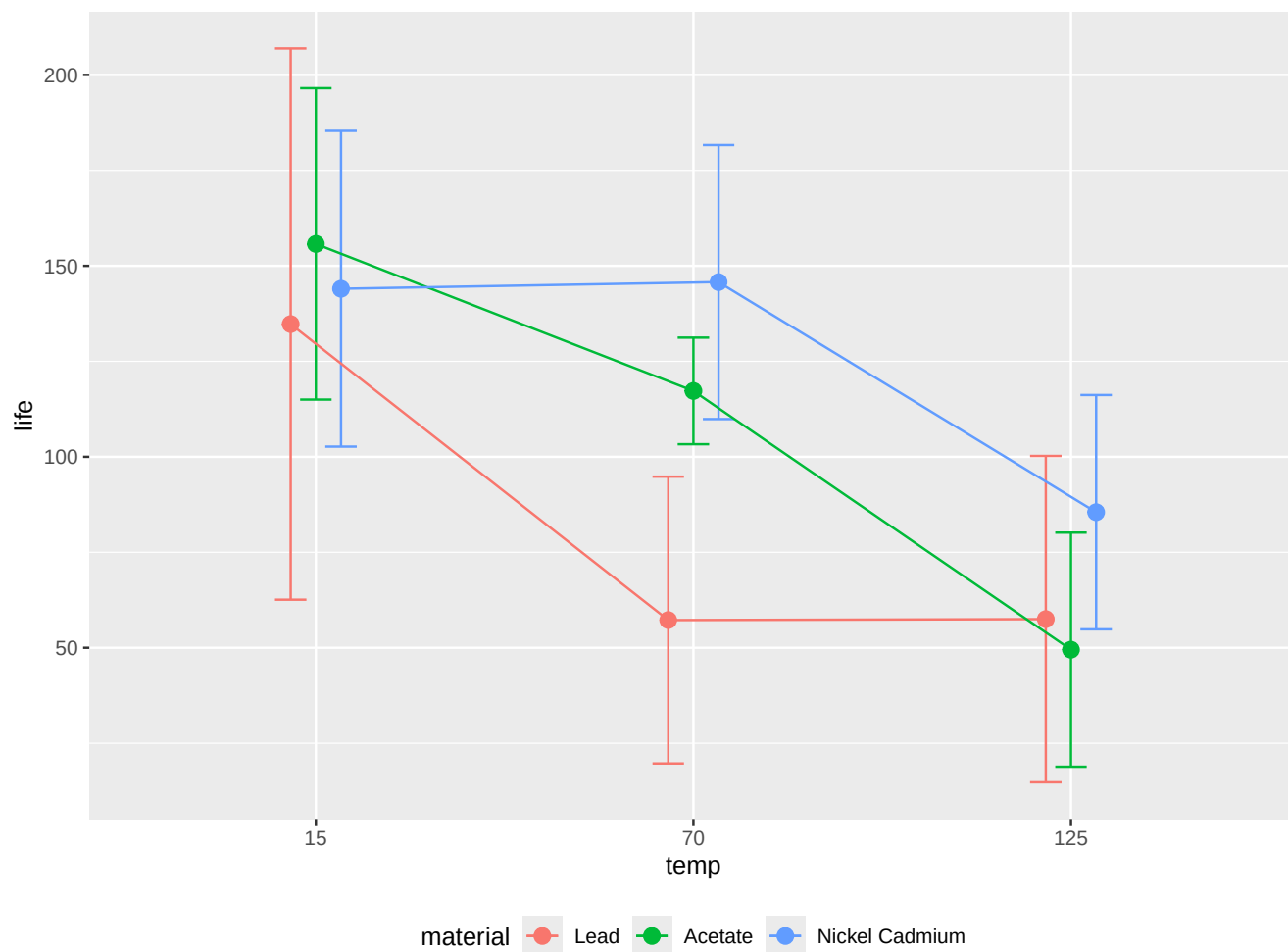


Figure 15: Means Plot for Batteries data

```

> pd <- position_dodge(0.2)
> ggplot(batsum,
+   aes(x=material, y=life, colour=temp, group=temp)) +
+   geom_errorbar(aes(ymin=life-ci, ymax=life+ci),
+     width=0.25, position=pd) +
+   geom_line(position=pd) +
+   geom_point(position=pd, size=3) +
+   theme(legend.position="bottom")

```

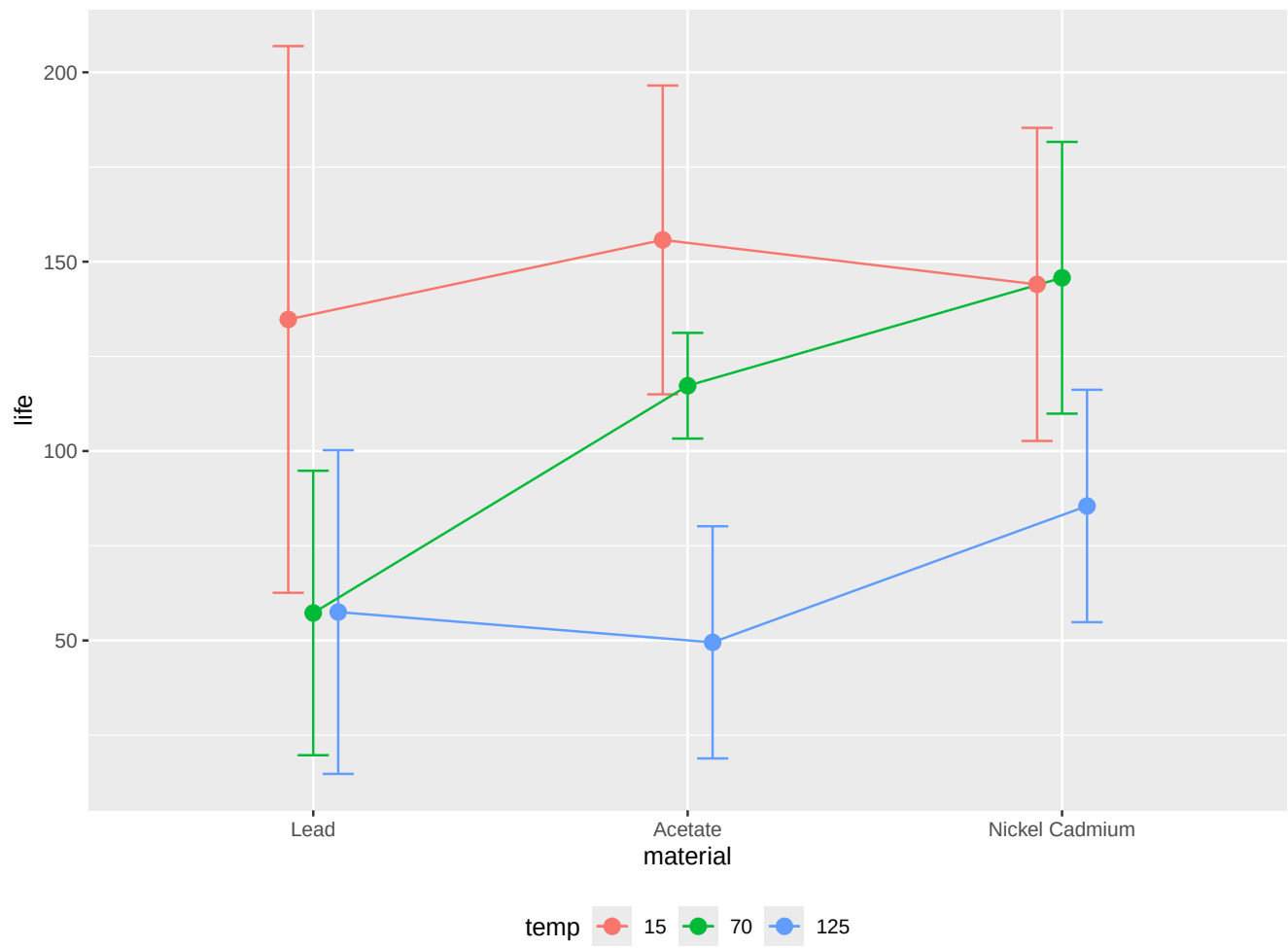


Figure 16: Means Plot for Batteries data