STAT 581 - Problem Set 7

Alex Towell (atowell@siue.edu)

Problem 1

A paired comparisons design is used to study the effect of machine operator on the measured running time (in secs.) of a fuse. A sample of $n = 10$ fuses is selected, and both operators provide a measurement of the running time on each of the selected fuses.

Part (a)

Test for a systematic difference in the measurements of the two operators. Compute the t_0 statistic, and the p-value. Provide an interpretation, stated in the context of the problem.

```
library("readxl")
data = read_excel("handout7data.xlsx")
A = as.factor(na.omit(data$o))
B = as.factor(na.omit(data$fuse))
y = na.omit(data$time)

op1 = na.omit(data$`operator 1`)
op2 = na.omit(data$`operator 2`)

paired.test(op1,op2)
```

```
##  sample.size mean.diff    sd.diff
##           10    -0.211 0.04094712
##  test statistic      p-value
##       -16.29518 5.484094e-08
##  estimated difference lower limit upper limit
##                -0.211       -0.24      -0.182
```

Let $D_i = Y_{1i} - Y_{2i}$. The test statistic in the paired design experiment is given by

$$t_0 = \frac{\bar{D}}{s_D/\sqrt{n}} = \frac{-0.211}{0.041/\sqrt{10}} = -16.3,$$

which under the reference distribution t_{n-1} has a p-value .000.

Interpretation

Θ

The experiment finds that the machine operator has an effect on the running time of fuses.

Part (b)

Compute the sample mean $\bar{D}$ and sample standard deviation s_D for the paired differences. Compute a 95% confidence interval estimate for the mean difference between operator measurements.

From the previous R output, we see that $\bar{D} = -0.211$ and $s_D = .041$. A confidence interval for $\bar{D}$ is given by

$$\mathrm{CI}(\bar{D}) = \bar{D} \pm t_{\alpha/2,n-1} \frac{s_D}{\sqrt{n}}.$$

Plugging in the known values (or just reading them off the previous output), we obtain

$$\begin{aligned} \mathrm{CI}(\bar{D}) &= -0.211 \pm 2.262 \frac{.041}{3.162} \\ &= [-.24, -.18]. \end{aligned}$$

As expected, $\mathrm{CI}(\bar{D})$ does not contain 0, and thus we would find the data incompatible with the hypothesis

$$H_0 : \tau_1 = \tau_2 = 0.$$

Part (c)

> Run the analysis as a randomized complete block design. Compute the F_0 statistic and the p-value. Explain how the statistical results are equivalent.

Let

$$\begin{aligned} A &:= \text{machine operator}, \\ B &:= \text{fuse}, \\ y &:= \text{running time}. \end{aligned}$$

We describe the observations in this experimental design by a random effects model using paired comparisons,

$$Y_{ijk} = \mu + \tau_i + \beta_j + \epsilon_{ijk} \begin{cases} i = 1, \ldots, a = 2 \\ j = 1, \ldots, b = 10 \\ k = 1, \ldots, n = 1 \end{cases}$$

where

$$\begin{aligned} \mu &= \text{overall main effect}, \\ \tau_i &= \text{effect of the } i\text{-th level of factor } A, \\ \beta_j &\sim \mathcal{N}(0, \sigma_\beta^2), \\ \epsilon_{ij} &\sim \mathcal{N}(0, \sigma^2). \end{aligned}$$

We compute the statistics with:

```
# factor A is the operator (fixed effect)
# factor B is the block factor (random effect)
# y is the response

# tau1 + tau2 = 0
contrasts(A)=contr.sum

# the 1 in front of B signifies that block levels are randomly selected from a
# common distribution.
random1.mod = lmer(y ~ (1|B) + A)

# the anova command is used to compute the test for fixed effects.
anova(random1.mod)
```

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
A	0.222605	0.222605	1	9	265.5328	1e-07

Explanation

The statistical results are equivalent since $F_0 = t_0^2$, i.e., $265.533 \approx (-16.3)^2$. Thus, we see that a paired difference analysis is equivalent to a block design analysis when $a = 2$.

Since we are not interested in estimating σ_β^2, we may also use the following R code:

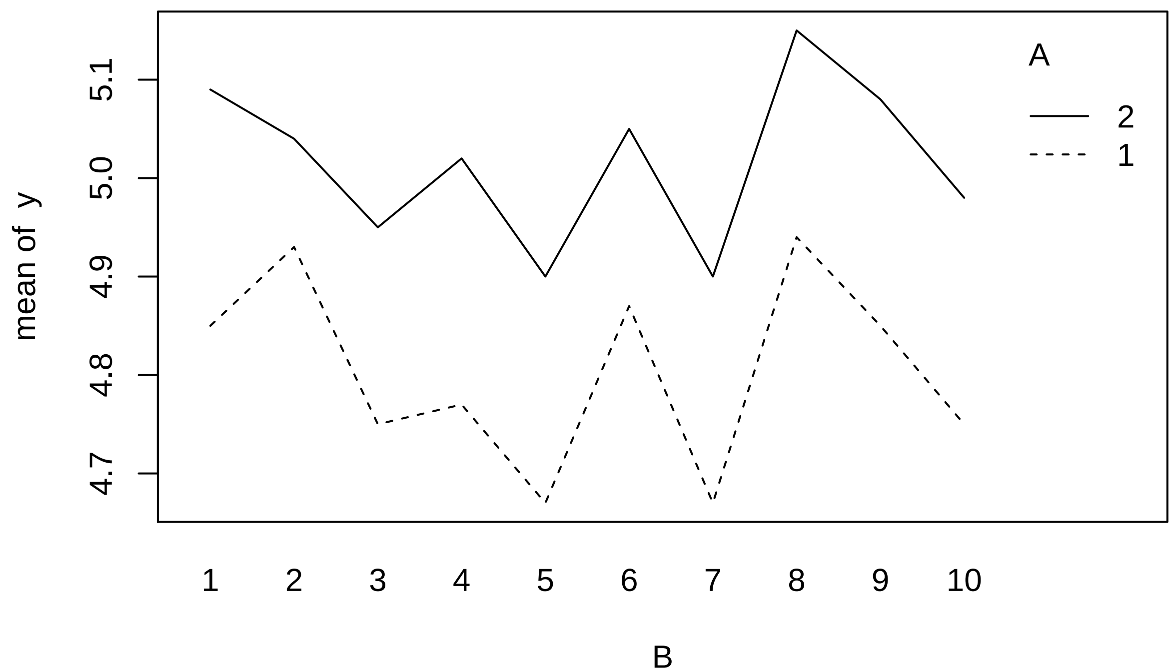
```
anova(aov(y~A+B))
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	0.222605	0.2226050	265.53280	1.00e-07
B	9	0.140145	0.0155717	18.57455	8.78e-05
Residuals	9	0.007545	0.0008383	NA	NA

Part (d)

Create an interaction plot to display the operator effect on measured running time.

```
interaction.plot(B,A,y)
```



We see that mean running time is consistently lower for operator 1 compared to operator 2 over the fuses in the sample, which gives us some confidence that we will be able to generalize to the larger population of fuses.

Problem 2

A randomized complete block design is used to study the effect of machine tip on the measured hardness (in Rockwell C-scale units) of a metal specimen. A sample of $b = 4$ metal specimens is randomly selected, and each of $a = 4$ machine tips produces a measurement on each of the selected metal specimens. The data is available on Blackboard as an Excel File.

Part (a)

Write the model and distributional assumptions for a randomized complete block design.

$$Y_{ij} = \mu + \tau_i + \beta_j + \epsilon_{ij} \begin{cases} i = 1, \dots, a \\ j = 1, \dots, b \end{cases}$$

where

$$\begin{aligned} \beta_j &\sim \mathcal{N}(0, \sigma_\beta^2) \\ \epsilon_{ij} &\sim \mathcal{N}(0, \sigma^2). \end{aligned}$$

Note that $\{\tau_i\}$ are fixed effects, $\{\beta_j\}$ are random effects, $\{\epsilon_{ij}\}$ are random errors, σ_β^2 is between block variance, and σ^2 is within block variance (variance in the response, or measurement variance).

Part (b)

Provide the algebraic formulas for MS_{tr}, MS_{bl}, and MS_{E}. State the expected value for each of the mean squares.

$$\begin{aligned} \text{MS}_{\text{tr}} &= \frac{b \sum_{i=1}^{a} (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2}{a-1}, \\ \text{MS}_{\text{bl}} &= \frac{a \sum_{i=1}^{b} (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2}{b-1}, \\ \text{MS}_{\text{E}} &= \frac{\sum_{i=1}^{a} \sum_{j=1}^{b} (Y_{ij} - \bar{Y}_{i\cdot} - \bar{Y}_{\cdot j} + \bar{Y}_{\cdot\cdot})^2}{(a-1)(b-1)}, \\ E(\text{MS}_{\text{tr}}) &= \frac{b \sum_{i=1}^{a} \tau_i^2}{a-1} + \sigma^2, \\ E(\text{MS}_{\text{bl}}) &= a\sigma_\beta^2 + \sigma^2, \\ E(\text{MS}_{\text{E}}) &= \sigma^2. \end{aligned}$$

Part (c)

Explain why it makes sense for an interaction effect to serve as a measure of error variance.

We are testing whether or not an observed treatment effect is generalizable to a larger populate; how the effect depends on the experimental units determines how well the effect can be generalized.

Main effects are aggregate effects. Since levels for a block (experimental unit) are not identifiable, we can only estimate the main effect of the treatment variable and we model the block level effects through a probability distribution.

Part (d)

Test for systematic differences in the measurements provided by the machine tips. Compute the F_0 statistic, and the p-value. Provide an interpretation, stated in the context of the problem.

Computation

```
tip = as.factor(data$tip) # tip level factor : fixed effect
spec = as.factor(data$specimen) # metal specimen : random effect
hard = data$hardness

# fixed effect tau1 + tau2 = 0 (for tip1 and tip2)
contrasts(tip)=contr.sum

# the 1 in front of s signifies that batch levels are randomly selected from a
# common distribution.
random.mod = lmer(hard ~ (1|spec) + tip)

#The anova command is used to compute the test for fixed effects.
anova(random.mod)
```

	Sum Sq	Mean Sq	NumDF	DenDF	F value	Pr(>F)
tip	0.385	0.1283333	3	9	14.4375	0.0008713

We see that $F_0 = \frac{\text{MS}_{\text{tr}}}{\text{MS}_{\text{E}}} = 14.438$ and p-value $= .001$.

Interpretation

The experiment finds that the tip factor has an effect on the hardness measurement.

Part (e)

> Perform pairwise comparisons using the Fisher LSD method to investigate differences between the machine tips. Provide the grouping information. Create box plots to display the tip effect on hardness measurement.

Fisher LSD comparison computations

```
comps = glht(random.mod,linfct = mcp(tip="Tukey"))
summary(comps,test=univariate())
```

```
##
##   Simultaneous Tests for General Linear Hypotheses
##
## Multiple Comparisons of Means: Tukey Contrasts
##
##
## Fit: lmer(formula = hard ~ (1 | spec) + tip)
##
## Linear Hypotheses:
##              Estimate Std. Error z value Pr(>|z|)
## 2 - 1 == 0  0.02500    0.06667   0.375   0.7077
## 3 - 1 == 0 -0.12500    0.06667  -1.875   0.0608 .
## 4 - 1 == 0  0.30000    0.06667   4.500 6.80e-06 ***
## 3 - 2 == 0 -0.15000    0.06667  -2.250   0.0244 *
## 4 - 2 == 0  0.27500    0.06667   4.125 3.71e-05 ***
## 4 - 3 == 0  0.42500    0.06667   6.375 1.83e-10 ***
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
## (Univariate p values reported)
```

```r
cld(summary(comps,test=univariate()))
```

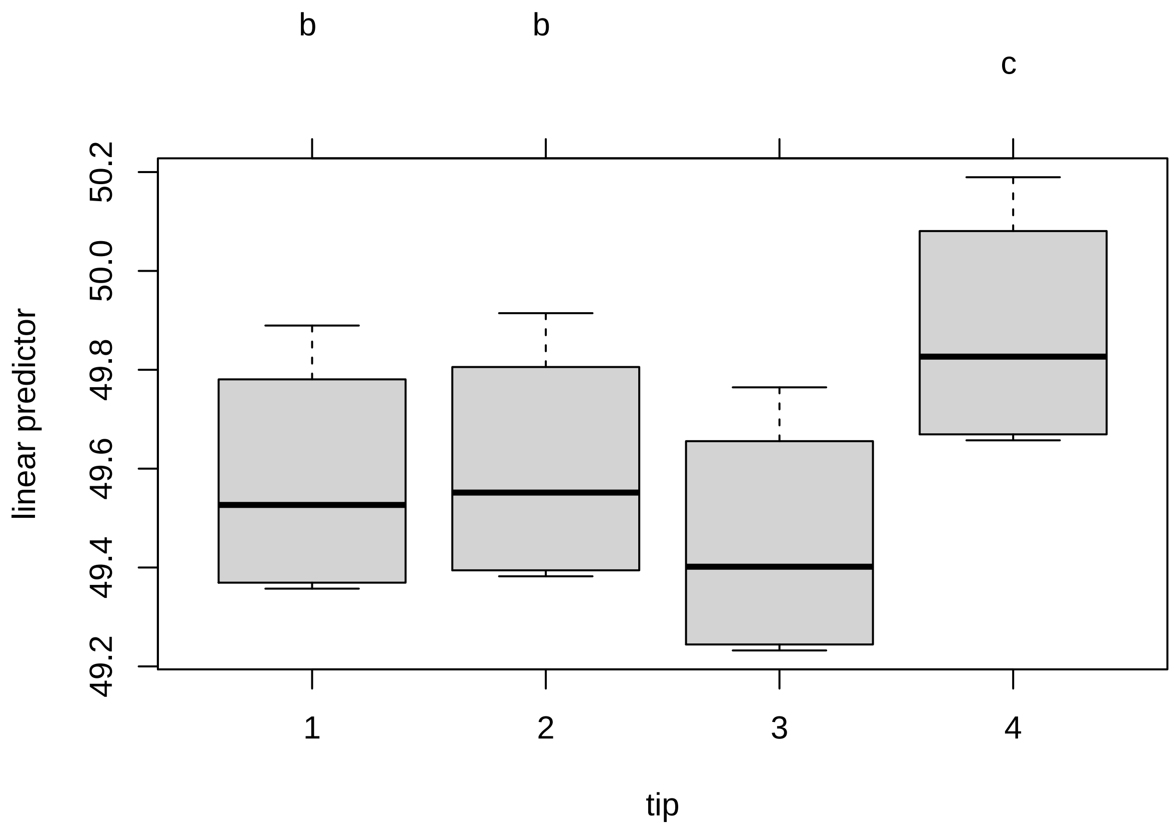
```
##    1    2    3    4
## "ab"  "b"  "a"  "c"
```

Grouping information table

tip			
1	A	B	
2		B	
3	A		
4			C

Boxplots

```r
plot(cld(summary(comps,test=univariate())))
```



Part (f)

> Compute estimates of the variance components. Explain when a block design is better than a completely randomized design.

Estimators of σ^2 (within-block variance) and σ^2_β (between block variance) are given by

$$\hat{\sigma}^2 = \mathrm{MS_E},$$
$$\hat{\sigma}^2_{bl} = \frac{\mathrm{MS_{bl}} - \mathrm{MS_E}}{a}.$$

The following outputs the standard deviation of the component estimates.

```
summary(random.mod)$varcor
```

```
##  Groups   Name        Std.Dev.
##  spec     (Intercept) 0.257930
##  Residual             0.094281
```

We see that $\hat{\sigma}^2_\beta = (.258)^2 = .067$ and $\hat{\sigma}^2 = (.094)^2 = 0.009$.

A block design is better than a CRD when between block variance is larger relative to within block variance.

Part (g)

Compute estimates of the fixed effect parameters.

i. Explain why block effects are modeled differently than treatment effects in this design.
ii. Explain what treatment effect is estimatable in this design.

```
# since the fixed effect estimates must sum to 0, the estimate at level a=4 is
# the negative of the sum of the fixed effect estimates at levels 1 through a-1.
estimates = summary(random.mod)
estimates.tau.hat = c(estimates$coefficients[1:4,1],0-sum(estimates$coefficients[2:4,1]))
names(estimates.tau.hat) = c("mu","tau1","tau2","tau3","tau4")
estimates.tau.hat
```

```
##     mu   tau1   tau2   tau3   tau4
## 49.625 -0.050 -0.025 -0.175  0.250
```

(i)

We consider the $b = 4$ specimens experimental units randomly selected from a large population of metal specimens. Since we have multiple measurements for each specimen (block), we may include a term for block effect. However, because the block levels have no identifiable features, we model block level effects through a probability distribution.

Note that a primary objective is to generalize to the larger population of *all* metal specimens. Indeed, going from the particular (observed sample) to the general (large population) is a unifying theme in all of statistics.

(ii)

The fixed effects, $\{\tau_i\}_{i=1}^4$, and the overall mean response μ, are the estimable effects. We do not estimate $\{\beta_j\}$, but only include it in the model as a noise reduction technique (reduction in sampling variance) when appropriate.

Appendix

Code

```
paired.test = function(y1,y2,alpha=.05)
{
```

```
  d = y1 - y2

  n = length(d)
  d.bar = mean(d)
  s.d = sd(d)
  SE = s.d/sqrt(n)

  t.0 = d.bar / SE
  p.value = 2*pt(abs(t.0),df=n-1,lower.tail = FALSE)

  t.mult = qt(alpha/2,lower.tail = FALSE, df=n-1)
  lower.est = d.bar - t.mult*SE
  upper.est = d.bar + t.mult*SE

  table1 = matrix(c(n,d.bar,s.d),nrow = 1)
  dimnames(table1) = list(c(""),c("sample.size","mean.diff","sd.diff"))
  print(table1)

  table2 = matrix(c(t.0,p.value),nrow = 1)
  dimnames(table2) = list(c(""),c("test statistic","p-value"))
  print(table2)

  table3 = matrix(c(d.bar,lower.est,upper.est),nrow = 1)
  dimnames(table3) = list(c(""),c("estimated difference","lower limit","upper limit"))
  print(table3,digits = 3)
}
```