Homework 9

ANOVA

Langdon Holmes

```
df <- data.frame(
  "neutral" = c(28.6, 28, 28),
  "a" = c(16.8, 23, 26.8),
  "b" = c(24.4, 16, 26.4),
  row.names = c("50ms", "100ms", "150ms")
)
```

	Area N	Area A	Area B	Row Means
50ms	28.6	16.8	24.4	23.27
100ms	28	23	16	22.33
150ms	28	26.8	26.4	27.07
Column Means	28.2	22.2	22.27	24.22

1 Test the hypothesis that the average response time across A and B is less than the neutral response time.

```
n=5
a=3
b=3
df_total <- n*a*b - a*b

ms_error <- 29.311
df_error <- 9-1

contrast <- function(aa, bb, ms_error, df_total, n_group, alpha=0.05, msg=""){

  aa <- as.numeric(aa)
  bb <- as.numeric(bb)
```

```r
  point_estimate <- mean(aa) - mean(bb)
  contrast_weights <- c(rep(1/length(aa), times=length(aa)),
                        rep(1/length(bb), times=length(bb))
  )

  standard_error <- sqrt(ms_error*sum(contrast_weights**2) / n_group)

  t <- point_estimate / standard_error
  p <- pt(t, df_total, lower.tail=FALSE)
  cv <- qt(alpha, df_total, lower.tail=FALSE)
  ci <- cv*standard_error

  cat(
    msg,
    "  \n",
    sprintf("$t = %1.3f,\\ p\\, (unadjusted) = %2.3f$,",
            round(t, 3),
            round(p, 3)
            ),
    sprintf("$%1.3f \\pm %2.3f$  \n",
            round(point_estimate, 3),
            round(ci, 3))
    )
}

contrast(colMeans(df)[1], colMeans(df)[2:3], ms_error, df_total, n*b)
```

$t = 3.485,\ p\,(unadjusted) = 0.001$, 5.967 ± 2.890

To test the null hypothesis that the average time-to-step in A and B is less than or equal to the time-to-step in N, I ran a one-sided t-test. The test was significant, which implies that the average time-to-step time in A and B is lower in than in the neutral group. In terms of the research problem, this means that the stimulus reduced time-to-step in at least one location x time-delay pair.

2 Suppose for each time lag you would like to test the difference between Area N and Area A and the difference between Area N and Area B, both in one-sided tests. Which direction should be specified in the alternative hypotheses of these tests in order to determine the effective brain area at each time lag? Obtain t statistics, unadjusted one-sided p-values and use Hochberg method to draw conclusion. There should be a total of six tests. You don't need to write the hypotheses.

```
contrast(df[1, 1], df[1, 2], ms_error, df_total, n, msg="Neutral vs. Area A  (50ms):")
```

Neutral vs. Area A (50ms):
$t = 3.446,\ p\,(unadjusted) = 0.001,\ 11.800 \pm 5.781$

```
contrast(df[1, 1], df[1, 3], ms_error, df_total, n, msg="Neutral vs. Area B  (50ms):")
```

Neutral vs. Area B (50ms):
$t = 1.227,\ p\,(unadjusted) = 0.114,\ 4.200 \pm 5.781$

```
contrast(df[2, 1], df[2, 2], ms_error, df_total, n, msg="Neutral vs. Area A (100ms):")
```

Neutral vs. Area A (100ms):
$t = 1.460,\ p\,(unadjusted) = 0.076,\ 5.000 \pm 5.781$

```
contrast(df[2, 1], df[2, 3], ms_error, df_total, n, msg="Neutral vs. Area B (100ms):")
```

Neutral vs. Area B (100ms):
$t = 3.505,\ p\,(unadjusted) = 0.001,\ 12.000 \pm 5.781$

```
contrast(df[3, 1], df[3, 2], ms_error, df_total, n, msg="Neutral vs. Area A (150ms):")
```

Neutral vs. Area A (150ms):
$t = 0.350,\ p\,(unadjusted) = 0.364,\ 1.200 \pm 5.781$

```
contrast(df[3, 1], df[3, 3], ms_error, df_total, n, msg="Neutral vs. Area B (150ms):")
```

Neutral vs. Area B (150ms):
$t = 0.467,\ p\,(unadjusted) = 0.322,\ 1.600 \pm 5.781$

The unadjusted p-values presented above do not correct for multiple comparisons. To determine significance, I will control for family-wise error using the Hochberg method. I can see that 4 unadjusted p-values are not significant (greater than 0.05) even before beginning the procedure, so we can retain these without any calculations. There are 2 remaining contrasts, so we choose the largest p-value and test at $\frac{\alpha}{2} = 0.025$, where 2 is the number of contrasts that have not yet been retained. Since 0.001 is less than 0.025, we reject the null for this contrast and also reject the smaller p-value. Note: I do not know which p-value is smaller, and I did not drill down because it would not affect the outcome.

To summarize: The difference between Neutral and Area A is significant only at a 50ms delay (time-to-step was lower for Area A). The difference between Neutral and Area B is significant only at a 100ms delay (time-to-step was lower for Area B).

3 Write a contrast that describes how much the difference between Areas A and B changes from time lag 100ms to time lag 50ms. Obtain an unadjusted 95% CI for this contrast.

```
contrast(
  diff(as.numeric(df[1, 2:3])),
  diff(as.numeric(df[2, 2:3])),
  ms_error,
  df_total,
  n,
  alpha=0.05/2,
  msg="$\\phi_{12 \\times 23}:$")
```

$\phi_{12 \times 23}:$
$t = 4.264,\ p\,(unadjusted) = 0.000,\ 14.600 \pm 6.944$

The contrast between areas A and B with time lags 50ms and 100ms was significant. This means that the effect of the time lag on response time is different for areas A and B. It seems that we are interested in a difference in either direction, so I divided our nominal alpha-value by 2.

4 If you wish to protect the family of all interaction contrasts, what is the CV?

```
k <- a*b # 9 cells
sqrt((k-1) * qf(0.05, k-1, df_total, lower.tail=FALSE))
```

```
[1] 4.203349
```

The family of all interaction contrasts is protected by Scheffe's method. The critical value is 4.203.

5 How many interaction contrasts are interactions of two pairwise comparisons? To obtain simultaneous CIs for this family of contrasts, what is the CV using Bonferroni method?

```
qf(0.05/9, k-1, df_total, lower.tail=FALSE)
```

```
[1] 3.367583
```

There are 9 pairwise interaction contrasts. To protect family-wise error rate in this family using Bonferroni's method, I divided our nominal alpha value by 9 to find the critical value of 3.368.

6 For the test in Problem 1 to achieve a power of 0.8 in detecting d = 0.5, what is the minimum per cell sample size? Attach screen shot of WebPower. Note if your effect size is specified as a positive number, the type of analysis should be "greater than".

```
d = 0.5
f <- d / sqrt(k*sum(c(1/1, 1/2, 1/2)**2))
wp.anova(k=a*b, n=NULL, f=f, power=0.8, type="greater")
```

```
Power for One-way ANOVA

    k        n         f alpha power
    9 335.2483 0.1360828  0.05   0.8

NOTE: n is the total sample size (contrast, greater)
URL: http://psychstat.org/anova
```

WebPower needs effect size in terms of f, which is related to d by $f = \frac{d}{\sqrt{k \sum_i c_i^2}}$. To achieve a power of 0.8 in detecting this contrast, one would need an overall sample size of 336, or a per-cell sample size of 38, assuming a balanced design.