Statistical tests for AA and ICP-MS

```r
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.0 --
```

```
## v ggplot2 3.3.2     v purrr   0.3.4
## v tibble  3.0.4     v dplyr   1.0.2
## v tidyr   1.1.2     v stringr 1.4.0
## v readr   1.4.0     v forcats 0.5.0
```

```
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```r
library(readr)
```

```r
aa_conc <- read_csv("~/chem313/313_icpms/data/sample_aa.csv")

icpms_conc <- read_csv("~/chem313/313_icpms/data/sample_icpms.csv")

icpms_all <- read_csv("~/chem313/313_icpms/data/sample_conc_allmetals.csv")
```

```r
icpms_conc_ppm <- icpms_conc %>%
  mutate(conc_mean = conc_mean/1000,
         conc_sd = conc_sd/1000,
         lower_ci = lower/1000,
         upper_ci = upper/1000) %>%
  select(-mass_frac,
         -lower,
         -upper)
```

```r
chromium <- rbind(icpms_conc_ppm, aa_conc) %>%
  mutate(n = case_when(
    site == "A" ~ 4,
    site == "B" ~ 5,
    site == "C" ~ 5,
    site == "D" ~ 3,
    site == "E" ~ 2,
    site == "F" ~ 3
  ))
```

```r
#inputs: means, sd, and n for both samples
#outputs: t value
ttest <- function(mean1, mean2, sd1, sd2, n1, n2){
  s_pooled <- sqrt((((sd1^2 * (n1 - 1)) + (sd2^2 *(n2-1)) ) / (n1 + n2 -2))
  t <- (abs(mean1 - mean2) / s_pooled) * sqrt(((n1 * n2) / (n1 + n2)))
  return(t)
}
```

```r
#t test between methods to make sure the instruments agree
ttest_a <- ttest(23.7337, 23.1362, 7.605, 7.347, 4, 4)
```

```r
ttest_b <- ttest(16.913, 16.206, 5.061, 5.182, 5, 5)
ttest_c <- ttest(34.009, 26.708, 16.942, 13.305, 5, 5)
ttest_d <- ttest(12.698, 9.873, 4.466, 4.116, 3, 3)
ttest_e <- ttest(19.492, 13.387, 8.835, 7.154, 2, 2)
ttest_f <- ttest(11.555, 6.970, 7.996, 5.152, 3, 3)

#t test between site a and others for AA
aa_ttest_ab <- ttest(23.162, 16.206, 7.347, 5.182, 4, 5)
aa_ttest_ac <- ttest(23.162, 26.708, 7.347, 13.205, 4, 5)
aa_ttest_ad <- ttest(23.162, 9.873, 7.347, 4.116, 4, 3) #significantly diff
aa_ttest_ae <- ttest(23.162, 13.387, 7.347, 7.154, 4, 2)
aa_ttest_af <- ttest(23.162, 6.970, 7.347, 5.152, 4, 3) #significantly diff

#t test between site b and others
aa_ttest_bc <- ttest(16.206, 26.708, 5.182, 13.205, 5, 5)
aa_ttest_bd <- ttest(16.206, 9.873, 5.182, 4.116, 5, 3)
aa_ttest_be <- ttest(16.206, 13.287, 5.182, 7.154, 5, 2)
aa_ttest_bf <- ttest(16.206, 6.970, 5.182, 5.152, 5, 3)

#t test between site c and others
aa_ttest_cd <- ttest(26.708, 9.873, 13.205, 4.116, 5, 3)
aa_ttest_ce <- ttest(26.708, 13.287, 13.205, 7.154, 5, 2)
aa_ttest_cf <- ttest(26.708, 6.970, 13.205, 5.152, 5, 3)

#t test between site d and others
aa_ttest_de <- ttest(9.873, 13.287, 4.116, 7.154, 3, 2)
aa_ttest_df <- ttest(9.873, 6.970, 4.116, 5.152, 3, 3)

#t test between site e and f
aa_ttest_ef <- ttest(13.287, 6.970, 7.154, 5.152, 2, 3)

#t test between site a and others for ICPMS
ttest_ab <- ttest(23.7337, 16.913, 7.605, 5.061, 4, 5)
ttest_ac <- ttest(23.7337, 34.009, 7.605, 16.942, 4, 5)
ttest_ad <- ttest(23.7337, 12.698, 7.605, 4.466, 4, 3)
ttest_ae <- ttest(23.7337, 19.492, 7.605, 8.835, 4, 2)
ttest_af <- ttest(23.7337, 11.555, 7.605, 7.996, 4, 3)

#t test betweeen site b and others
ttest_bc <- ttest(16.913, 34.009, 5.061, 16.942, 5, 5)
ttest_bd <- ttest(16.913, 12.698, 5.061, 4.466, 5, 3)
ttest_be <- ttest(16.913, 19.492, 5.061, 8.835, 5, 2)
ttest_bf <- ttest(16.913, 11.555, 5.061, 7.996, 5, 3)

#t test between site c and others
ttest_cd <- ttest(34.009, 12.698, 16.942, 4.466, 5, 3)
ttest_ce <- ttest(34.009, 19.492, 16.942, 8.835, 5, 2)
ttest_cf <- ttest(34.009, 11.555, 16.942, 7.996, 5, 3)

#t test between site d and others
ttest_de <- ttest(12.698, 19.492, 4.466, 8.835, 3, 2)
ttest_df <- ttest(12.698, 11.555, 4.466, 7.996, 3, 3)

#t test between site e and f
```

```
ttest_ef <- ttest(19.492, 11.555, 8.835, 7.996, 2, 3)
```

```
#anova test for rest of icpms metals
icpms_aov <- aov(conc_mean ~ site, data = icpms_all)
summary(icpms_aov)
```

```
##              Df    Sum Sq   Mean Sq F value Pr(>F)
## site          5 1.473e+10 2.945e+09    0.26  0.931
## Residuals    30 3.401e+11 1.134e+10
```