# 17: TWO SAMPLE CONFIDENCE INTERVALS

Stat250 S25
Prof Amanda Luby

---

Goal: develop an **interval estimate** for a difference in population means: $\mu_1 - \mu_2$

**Example:** There are two types of forensic analyses that are typically performed in cases where firearms evidence has been collected: bullets and cartridges. We are interested in whether there is a difference in the average *inconclusive rate* (how often the firearms examiner cannot come to a definitive conclusion) for these two evidence types. Some results from a study are included below:



| type | mean | sd | n |
|------|------|------|-----|
| bullet | 0.547 | 0.276 | 38 |
| cartridge | 0.461 | 0.223 | 40 |

Let's review our confidence interval toolkit so far:

1. Construct a bootstrap distribution and find a **percentile-based** confidence interval

2. Either (a) assume a normal population or (b) use the CLT to construct a **formula t-based** confidence interval

3. Assume a different population distribution and construct a **pivot-based** confidence interval

4. Use the bootstrap to find the distribution of the T statistics to construct a **bootstrap t** confidence interval

**Bootstrap percentile interval**

```
x <- bullets_inc
y <- cc_inc

n1 = length(x)
n2 = length(y)
N = 10^4
boot_diff = numeric(N)

for(i in 1:N){
  x_boot <- sample(x, size = n1, replace = TRUE)
```

```
  y_boot <- sample(y, size = n2, replace = TRUE)
  boot_diff[i] <- mean(x_boot, na.rm = TRUE) - mean(y_boot, na.rm = TRUE)
}

quantile(boot_diff, probs = c(.025, .975))
```

```
       2.5%       97.5%
-0.02295906  0.19685965
```

## CLT-based t confidence interval

```
qt(.975, df = 37)
```

```
[1] 2.026192
```

```
t.test(bullets_inc, cc_inc, conf.level = .95)$conf
```

```
[1] -0.02682935  0.20045508
attr(,"conf.level")
[1] 0.95
```

## Bootstrap t confidence interval

```
n1 = length(bullets_inc)
n2 = length(cc_inc)
N = 10^4
boot_tstar <- numeric(N)
for(i in 1:N){
  x_boot <- sample(x, size = n1, replace = TRUE)
  y_boot <- sample(y, size = n2, replace = TRUE)
  boot_tstar[i] <- ((mean(x_boot) - mean(y_boot)) - (mean(x) - mean(y)))/sqrt(var(x_boot)/n1 + var(y_boot
}

quantile(boot_tstar, probs = c(.025, .975))
```
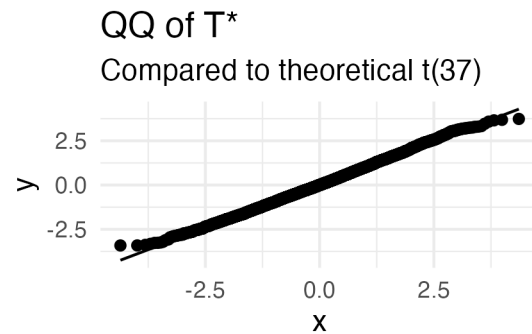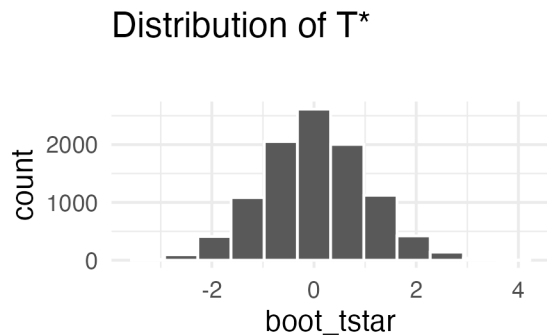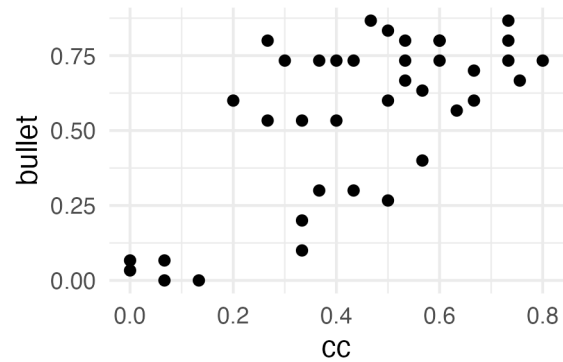
```
      2.5%        97.5%
-1.929443   2.097633
```

```
(mean(x)- mean(y)) - quantile(boot_tstar, c(.975, .025))*sqrt(var(x)/n1 + var(y)/n2)
```

```
      97.5%            2.5%
-0.03274359   0.19678315
```

## Distribution of T*



## QQ of T*
### Compared to theoretical t(37)



**Watch out!**

**Example**: Oops! I forgot to tell you that this study actually had the same examiners in both samples. (There are just two missing values in `bullets_inc`).

3

```r
x <- bullets_inc - cc_inc[-c(1,2)]
n = length(x)
N = 10^4
boot_diff = numeric(N)

for(i in 1:N){
  x_boot <- sample(x, size = n, replace = TRUE)
  boot_diff[i] <- mean(x_boot, na.rm = TRUE)
}

quantile(boot_diff, probs = c(.025, .975))
```

```
      2.5%       97.5%
0.03888889 0.16491228
```

```r
t.test(bullets_inc, cc_inc[-c(1,2)], paired = TRUE)$conf
```

```
[1] 0.03670524 0.16738833
attr(,"conf.level")
[1] 0.95
```

```r
t.test(bullets_inc - cc_inc[-c(1,2)])$conf
```

```
[1] 0.03670524 0.16738833
attr(,"conf.level")
[1] 0.95
```