Chapter 3

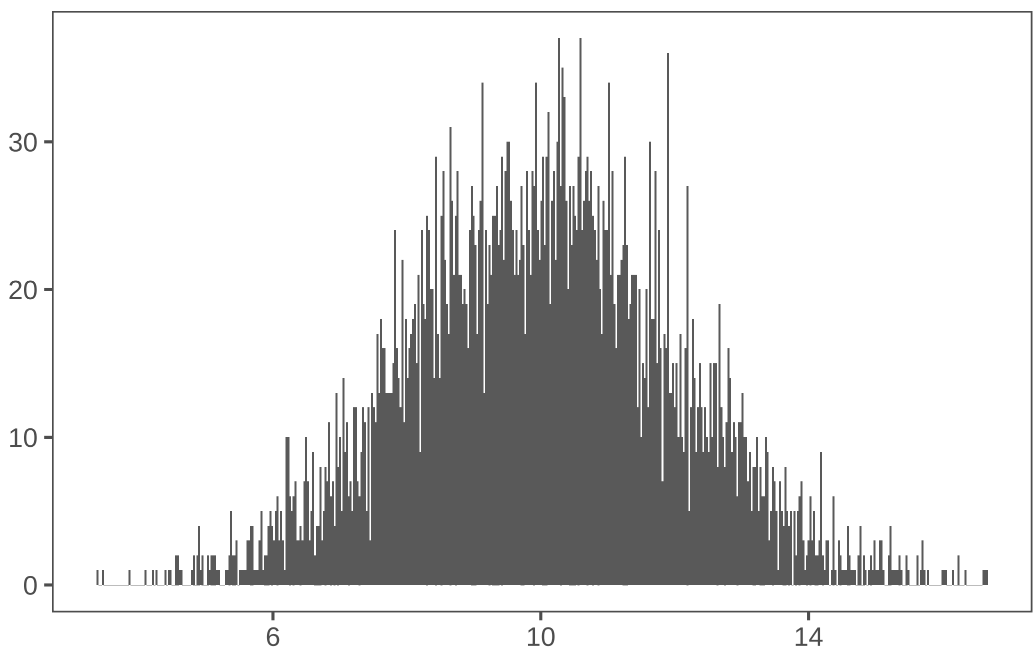
```r
library(tidyverse)
library(knitr)
```

Our effect size stuff:

$$\hat{\theta}_k = \theta_k + \epsilon_k$$

```r
set.seed(420)
sample <- rnorm(n = 5000, mean = 10, sd = 2)

as_tibble(sample) |>
  ggplot(aes(value)) +
  geom_histogram(bins = 500) +
  labs(y = NULL, x = NULL) +
  ggthemes::theme_few()
```



$$SE = \frac{s}{\sqrt{n}}$$

Therefore: N gets larger -> SE gets smaller -> study gets more precise

```
sample |>
  as_tibble() |>
  summarise(
    mean = mean(value),
    sd = sd(value),
    se = sd(value)/sqrt(500)) |>
  kable(digits = 3)
```

mean	sd	se
10.028	1.998	0.089

Proportions

How many units fall into a (sub)-group. Ranges from 0 to 1.

$$p = \frac{k}{n}$$

$$SE_p = \sqrt{\frac{p(1-p)}{n}}$$

```
tibble(
  n = 300,
  k = 150
) |>
  summarise(
    p = k / n,
    SE_p = sqrt((p * (1 - p)) / n)
  )
```

```
# A tibble: 1 x 2
      p   SE_p
  <dbl>  <dbl>
1   0.5 0.0289
```

Correlation

Expresses the amount of covariation between two variables

Pearsons Product-Moment-Correlation

Ranges from -1 to 1

$$r_{xy} = \frac{\sigma^2_{xy}}{\sigma_x \sigma_y}$$

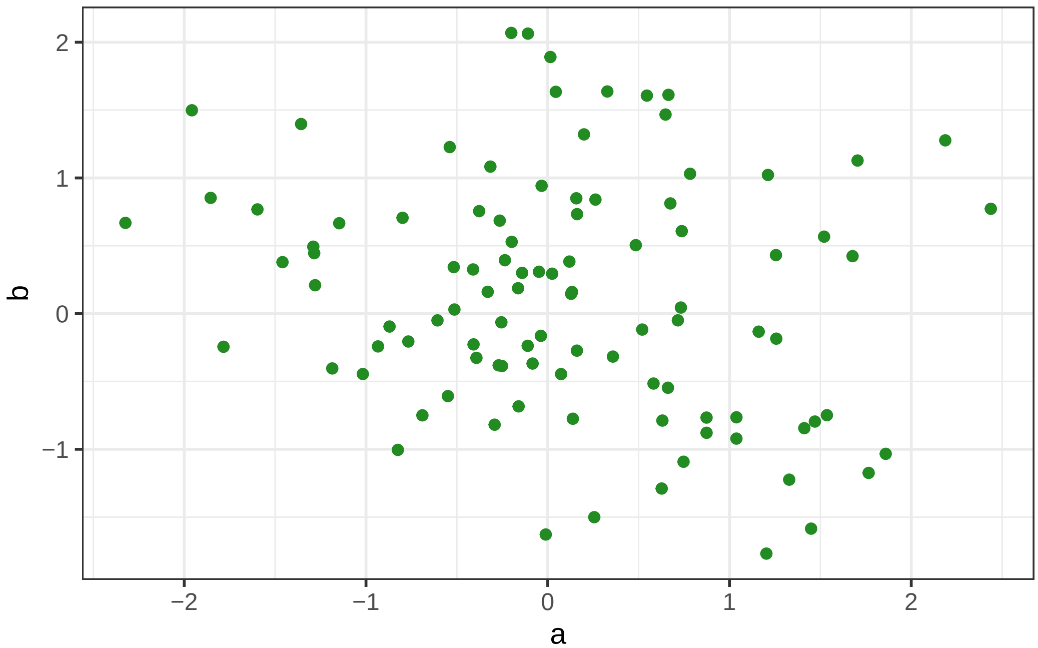
$$SE_{r_{xy}} = \frac{1 - r^2_{xy}}{\sqrt{n-2}} \tag{3.10}$$

```
set.seed(1999)
corr <- tibble(
  a = rnorm(100),
  b = rnorm(100)
)

cor(corr$a,corr$b) |>
kable(digits = 2)
```

x
-0.2

```
corr |>
  ggplot(aes(a,b)) +
  geom_point(col = "forestgreen") +
  theme_bw()
```



2
1
0
−1
b
−2
−1
0
1
2
a