**Key points**

* We can use statistical theory to compute the probability that a given interval contains the true parameter p.
* 95% confidence intervals are intervals constructed to have a 95% chance of including p. The margin of error is approximately a 95% confidence interval.
* The start and end of these confidence intervals are random variables.
* To calculate any size confidence interval, we need to calculate the value z for which Pr(−z≤Z≤z) equals the desired confidence. For example, a 99% confidence interval requires calculating z for Pr(−z≤Z≤z)=0.99.
* For a confidence interval of size q, we solve for z=1−1−q2.
* To determine a 95% confidence interval, use z <- qnorm(0.975). This value is slightly smaller than 2 times the standard error.

**Code: geom\_smooth confidence interval example**

The shaded area around the curve is related to the concept of confidence intervals.

data("nhtemp")

data.frame(year = as.numeric(time(nhtemp)), temperature = as.numeric(nhtemp)) %>%

ggplot(aes(year, temperature)) +

geom\_point() +

geom\_smooth() +

ggtitle("Average Yearly Temperatures in New Haven")

**Code: Monte Carlo simulation of confidence intervals**

Note that to compute the exact 95% confidence interval, we would use qnorm(.975)\*SE\_hat instead of 2\*SE\_hat.

p <- 0.45

N <- 1000

X <- sample(c(0,1), size = N, replace = TRUE, prob = c(1-p, p)) # generate N observations

X\_hat <- mean(X) # calculate X\_hat

SE\_hat <- sqrt(X\_hat\*(1-X\_hat)/N) # calculate SE\_hat, SE of the mean of N observations

c(X\_hat - 2\*SE\_hat, X\_hat + 2\*SE\_hat) # build interval of 2\*SE above and below mean

**Code: Solving for**z**with qnorm**

z <- qnorm(0.995) # calculate z to solve for 99% confidence interval

pnorm(qnorm(0.995)) # demonstrating that qnorm gives the z value for a given probability

pnorm(qnorm(1-0.995)) # demonstrating symmetry of 1-qnorm

pnorm(z) - pnorm(-z) # demonstrating that this z value gives correct probability for interval