AGH Practicals 10 April 2021

Background:

Suppose one observes y that is Binomially distributed with sample size n and probability of success p. The standard 90% confidence interval for p is given by

$$C(y) = \left(\hat{p} - z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}, \hat{p} + z_{\alpha/2}\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}\right)$$

where $\hat{p} = y/n$. We use this procedure under the assumption that

$$P(p \in C(y)) = 0.90 \quad \text{for all } 0 < p < 1.$$

True probability of success is 0.5:

Suppose that samples of size $n = 20$ are taken and the true value of the proportion is $p = 0.50$. Simulate a value of y and compute the 90% confidence interval. Repeat this a total of 20 times and estimate the true probability of coverage $P(p \in C(y))$.

```
#Setting seed number
set.seed(12419)

#Compiler of coverage probability
covprob <- c()

for(i in 1:20) {

#Estimate of p
p_hat <- rbinom(n = 1, size = 20, prob = 0.5)/20

#90% CI lower bound
CI_low <- p_hat - 1.645*sqrt(p_hat*(1-p_hat)/20)

#90% CI upper bound
CI_up <- p_hat + 1.645*sqrt(p_hat*(1-p_hat)/20)

#Compiling coverages
covprob <- c(covprob,ifelse(CI_low <= 0.5 & CI_up >= 0.5,1,0))

}

#Estimate of coverage probability
mean(covprob)
```

```
## [1] 0.85
```

Another Solution:

Suppose that samples of size $n = 20$ are taken and the true value of the proportion is $p = 0.50$. Simulate a value of y and compute the 90% confidence interval. Repeat this a total of 20 times and estimate the true probability of coverage $P(p \in C(y))$.

```
# Binomial Confidence Interval
bin.ci <- function(N, n, p, alpha) {
  y <- rbinom(N, size = n, prob = p)
  p.hat <- y/n
  z <- qnorm(1 - (alpha/2))
  se <- sqrt(p.hat*(1 - p.hat)/n)
  ci.low <- p.hat - z*se
  ci.up <- p.hat + z*se
  return(as.numeric(ci.low <= p)*as.numeric(ci.up >= p))
}

# Monte Carlo Simulation
set.seed(12419)
ci.simulation <- function() {
  sum(bin.ci(N = 20, n = 20, p = 0.50, alpha = 0.10))/20
}


ci.vector <- replicate(10000, ci.simulation())
# Answer
print(ci.vector[1])
```

```
## [1] 0.85
```

```
# Something Extra
print(ci.vector[1:10])
```

```
##  [1] 0.85 0.95 0.75 0.90 0.80 0.85 0.90 0.90 0.80 0.95
```

```
print(c(mean(ci.vector), sd(ci.vector)))
```

```
## [1] 0.88484000 0.07039297
```

True probability of success is 0.05:

Suppose that $n = 20$ and the true value of the proportion is $p = 0.05$. Simulate 20 binomial random variates with $n = 20, p = 0.05$, and for each simulated y, compute a 90% confidence interval. Estimate the true probability of coverage.

```r
#Estimates of p
p_hat <- rbinom(n = 20, size = 20, prob = 0.05)/20

#90% CI lower bounds
CI_low <- p_hat - 1.645*sqrt(p_hat*(1 - p_hat)/20)

#90% CI upper bounds
CI_up <- p_hat + 1.645*sqrt(p_hat*(1 - p_hat)/20)

#Estimate of coverage probability
mean(ifelse(CI_low <= 0.05 & CI_up >= 0.05,1,0))
```

```
## [1] 0.6
```