Constructing $(1-\alpha)100\%$ confidence interval for a population proportion in R

Suppose that we have a sample of size $n$ consists of binary responses (0 or 1). A $(1-\alpha)100\%$ confidence interval for a population proportion of 1's ($= p$) in the sample is given by

$$\left[\hat{p} - z_{\alpha/2} \text{ standard error}, \hat{p} + z_{\alpha/2} \text{ standard error}\right]$$

where $\hat{p}$ is the sample proportion of 1's, standard error measures the variation of the sample proportion and equals to $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ and $z_{\alpha/2}$ is the corresponding z-score for the $(1-\alpha)100\%$ confidence interval.

To motivate the use of R for constructing C.I.'s for $p$, consider the following sample of size $n = 100$ simulated from a bernoulli distribution with probability of success $p$, which is unknown.

```
> set.seed(390)
> n = 100
> x = sample(c(0,1), n , replace=TRUE)
> x
```

```
  [1] 1 0 1 0 0 1 0 0 1 1 1 0 1 1 1 1 0 0 0 1 0 1 0 1 0 0 0 0 1 1 1 1 1 0 1 0 0
 [38] 1 1 1 1 0 0 0 1 0 0 1 0 1 0 1 1 0 0 0 1 1 1 0 1 1 1 1 1 0 1 1 0 1 1 0 0 0
 [75] 1 0 0 0 1 1 1 0 1 1 0 0 0 1 0 0 0 0 0 1 0 1 0 0 1 0
```

The variable `x` is a vector of dimension 100 and of values either 0 or 1. The sample proportion of $1's$ is given by:

```
> phat = sum(x)/n
> phat
```

```
[1] 0.5
```

and the standard error is:

```
> se = sqrt(phat * ( 1 - phat ) / n)
> se
```

```
[1] 0.05
```

Suppose that $\alpha = 0.05$, it means that we want to construct a 95% confidence interval.

```
> alpha = 0.05
> z = qnorm(1-alpha/2)
> z
```

```
[1] 1.959964
```

the variable `z` is the z-score corresponding to the 95% confidence interval.

The $(1-\alpha)100\%$ confidence interval is given by:

```
> c(phat - z*se , phat + z*se)
```

```
[1] 0.4020018 0.5979982
```

Therefore, the 95% confidence interval for the population proportion $p$ is

$$[0.402, 0.598].$$

You may also want to verify that, using the same sample above, the 90% confidence interval for $p$ is $[0.4178, 0.5822]$. Note that the length of the confidence interval increases when the value of $\alpha$ increases.