Exercise 2 Solution

Question 1

1.1

```
p0<-0.8 #null hypthesis
n<-100 #sample size
(Z<-(85-n*p0)/sqrt(n*p0*(1-p0))) #Z statistic
```

```
## [1] 1.25
```

```
#rejection rule: reject if |Z|>97.5% quantile of N(0,1)
abs(Z)>qnorm(0.975) #not rejected
```

```
## [1] FALSE
```

Reject H_0 if $|Z| > z_{0.025} \approx 1.96$, i.e. $X > np_0 + z_{0.025}\sqrt{np_0(1-p_0)}$ or $X < np_0 - z_{0.025}\sqrt{np_0(1-p_0)}$. These endpoints are

```
n*p0+c(-1,1)*qnorm(0.975)*sqrt(n*p0*(1-p0))
```

```
## [1] 72.16014 87.83986
```

So we reject H_0 if $X \leq 72$ or $X \geq 88$. 85 is not in the rejection region, so we don't reject H_0.

1.2

```
sum(dbinom(
    c(0:floor(n*p0-qnorm(0.975)*sqrt(n*p0*(1-p0))),
      ceiling(n*p0+qnorm(0.975)*sqrt(n*p0*(1-p0))):n),n,p0))
```

```
## [1] 0.05948038
```

It's close to the 0.05 from normal approximation.

1.3

```
sum(dbinom(
    c(0:floor(n*p0-qnorm(0.975)*sqrt(n*p0*(1-p0))),
      ceiling(n*p0+qnorm(0.975)*sqrt(n*p0*(1-p0))):n),n,0.9))
```

```
## [1] 0.8018215
```

1.4

```
phat<-85/100
SE<-sqrt(phat*(1-phat)/n)
phat+qnorm(c(0.025,0.975))*SE #95% CI
```

```
## [1] 0.7800153 0.9199847
```

The 95% CI covers the null proportion, so they give the same conclusions. The 95% CI further provides uncertainty of the sample proportion as an estimator of the population proportion.

Question 2

2.1

The null hypothesis is that the mean of test concentration value changes between the fresh sample and the aged sample is 0. The alternative hypothesis is that the mean of test concentration value changes between the fresh sample and the aged sample is not 0.

Let μ denote the mean of test concentration value changes between the fresh sample and the aged sample, then

$$H_0 : \mu = 0; H_1 : \mu \neq 0.$$

2.2

Let $\bar{X}$ be the sample mean difference. Then $Z = \frac{\bar{X}}{\sqrt{15/n}} = \bar{X}$ (here $n = 15$), and the rejection rule is $|Z| > z_{0.025} \approx 1.96$, that is $|\bar{X}| > z_{0.025}$.

```
data<-c(-5,-2,-1,-1,0,0,2,3,4,4,5,5,6,6,11)
n<-length(data)
xbar<-mean(data)
(Z<-xbar/(sqrt(15/n)))
```

```
## [1] 2.466667
```

```
abs(xbar)>qnorm(0.975)*sqrt(15/n) #reject null hypothesis?
```

```
## [1] TRUE
```

```
2*pnorm(abs(xbar),mean=0,sd=sqrt(15/n),lower.tail=FALSE) #p-value
```

```
## [1] 0.01363772
```

The null hypothesis that the mean of test concentration value changes between the fresh sample and the aged sample is 0 is rejected. The p-value is 0.014.

2.3

```
SE<-sqrt(15/length(data))
mean(data)+qnorm(c(0.025,0.975))*SE
```

```
## [1] 0.5067027 4.4266307
```

The 95% CI does not cover 0. The conclusion is the same.

2.4

```
N<-1e3 #number of samples
set.seed(3256)
n<-length(data)
```

```r
sim_result<-replicate(N,{
    sim_data<-rnorm(n,0,sd=sqrt(15))
    Z<-mean(sim_data)/sqrt(15/n)
    abs(Z)>qnorm(0.975)
})

mean(sim_result)
```

```
## [1] 0.049
```

```r
#Another example with non-normal distribution
sim_result<-replicate(N,{
    sim_data<-rpois(n,15)-15 #mean 0, variance 15
    Z<-mean(sim_data)/sqrt(15/n)
    abs(Z)>qnorm(0.975)
})

mean(sim_result)
```

```
## [1] 0.05
```

The type I error probability is very close to 0.05.