Stat 245: Question 7

Mufitcan Atalay

4/18/2018

a)

```
(1-pnorm(10, mean = 0, sd = 1))*2
```

```
## [1] 0
```

This is the tenth standard deviation so it is understandable that we get a p-value that is so small that `R` rounds it 0. b)

```
(1-pnorm(1, mean = 0, sd = 1))*2
```

```
## [1] 0.3173105
```

This makes sense since we are looking at the 1st standard deviation from the assumed mean of 0.

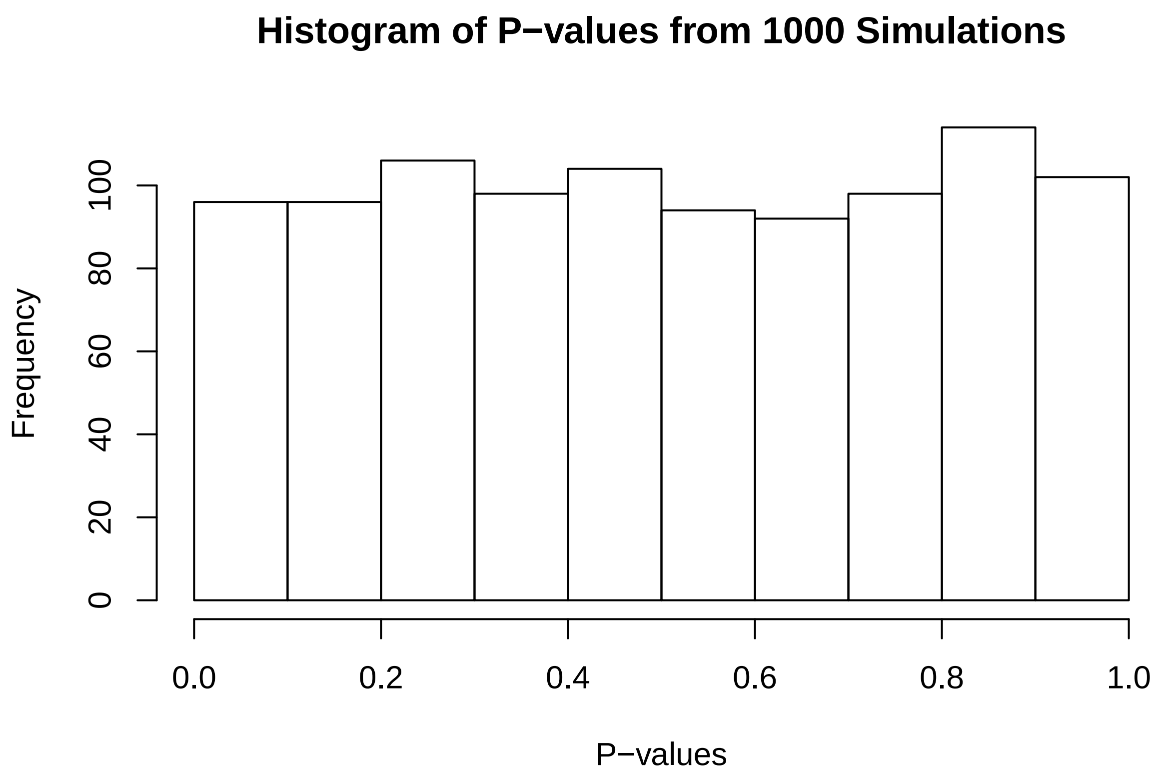
c)

```
rv1 <- rnorm(100, mean = 50, sd =1)
mrv1 <- mean(rv1)
z1 = abs((mrv1 - 50)/(1/sqrt(100)))
pval1 <- (1-pnorm(z1, mean = 0, sd = 1))*2
pval1
```

```
## [1] 0.104391
```

d)

```
nrSamples = 1000 #number of repetitions
rvs <- lapply(1:nrSamples, function(x) rnorm(n=100, mean = 50, sd = 1 )) # Creating 1000 100-sized samp
means <- sapply(rvs, mean) # Obtaining 1000 means from our samples
zs = abs((means - 50)/(1/sqrt(100)))
pvals <- (1-pnorm(zs, mean = 0, sd = 1, lower.tail = ))*2 # Obtainig 1000 p-values
hist(pvals, main = "Histogram of P-values from 1000 Simulations", xlab = "P-values", ylab = "Frequency")
```



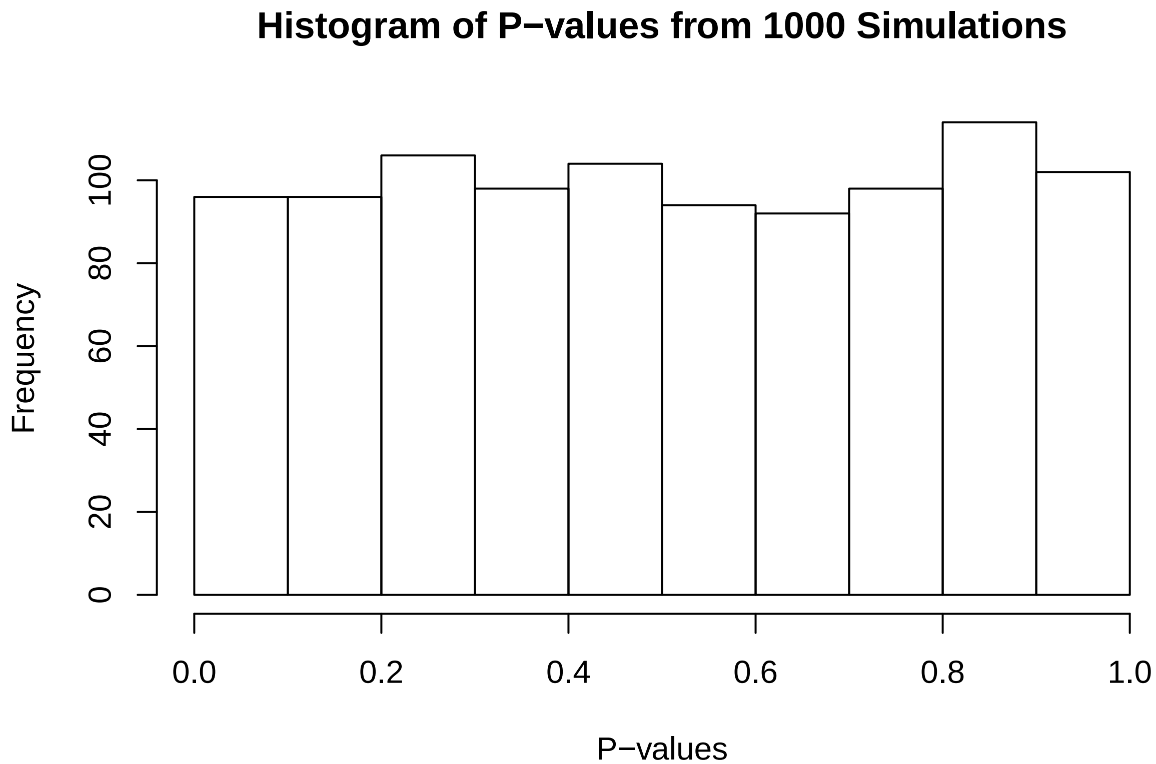
e)

```
rv2 <- rnorm(100, mean = 0, sd =1)
mrv2 <- mean(rv2)
z2 = abs((mrv2)/(1/sqrt(100)))
pval2 <- (1-pnorm(z2, mean = 0, sd = 1))*2
pval2
```

```
## [1] 0.007336542
```

f)

```
nrSamples = 1000 #number of repetitions
rvs1 <- lapply(1:nrSamples, function(x) rnorm(n=100, mean = 0, sd = 1 )) # Creating 1000 100-sized samp
means1 <- sapply(rvs1, mean) # Obtaining 1000 means from our samples
zs1  = abs((means1)/(1/sqrt(100)))
pvals1 <- (1-pnorm(zs1, mean = 0, sd = 1))*2 # Obtainig 1000 p-values
hist(pvals, main = "Histogram of P-values from 1000 Simulations", xlab = "P-values", ylab = "Frequency"
```



Finding the number of values below $\alpha = 0.05$

```
sum(pvals1 < 0.05)
```

```
## [1] 56
```

There are rougly 50 p-values below 0.05 (when I knit the document this number is randomly generated so it is impossible to really know the value, but it is always close to 50). This is because p-values form a uniform distribution when the null hypothesis is true, so the number of p-values below 0.05 should equal the probability $\alpha = 0.05$ times our size of a 1000. The whole point of a p-value is that we are able to take a z-test or any other form of test and turn it into a value which we can compare to our α, so it must follow a uniform distribution $U(0, 1)$. Since our the α we select is 0.05, there should only be 0.05 proportion out of all the p-values which are below 0.05, which is the case.

g) As stated above the point of p-values is that you are able to take any test statistics and compare it to a desired confidence level α. This means that p-values must follow a uniform distribution that is bound between 0 and 1, which is what we get from our answeres to d) and f). In this way the proportion of p-values below our desired level α will always be roughly aroound α.