

Assignment 11

AVIK BANERJEE (3374885), SOUMYADEEP BHATTACHARJEE (3375428)

Exercise 11.1

H_0 : Population mean is equal to 3.5 and H_1 : Population mean is greater than 3.5

- Integration performed using Simpson's rule.

```
x <- c(3.24,4.39,5.24,3.83,3.50,3.75,4.06,3.83,3.54,3.20,4.28,3.65,3.01,
      4.69,3.32)
sigma <- 0.76
mu <- 3.5
mux <- mean(x)
sigmax <- sigma/sqrt(length(x))
z <- (mux-mu)/sigmax
z
n=10000
arr <- seq(0, z, length=n)
fsum=0
i=0
for(x in arr){
fsum=fsum+(3-(-1)^i)*1/sqrt(2*pi)*exp(-x^2/2)
i=i+1
}
rs=fsum*((arr[2]-arr[1])/3)
r=rs*2
r1=1-r
p=r1/2
p
```

Outputs:

```
> z
[1] 1.708869
> p
[1] 0.04370433
```

Hence the p-value is 0.0437.

- The test here is one-tailed and $\alpha = 0.05$. Since $p = 0.0437 < \alpha$, we can reject the null hypothesis.

Exercise 11.2

H_0 : Population mean is equal to 3.5 and H_1 : Population mean is greater than 3.5

- t-test with integration using Simpson's rule. The t-distribution PDF is given by

$$\text{PDF}(x, \text{df}) = \frac{\Gamma\left(\frac{\text{df}+1}{2}\right)}{\sqrt{\pi \times \text{df}} \times \Gamma\left(\frac{\text{df}}{2}\right) \times \left(\frac{1+x^2}{\text{df}}\right)^{\frac{\text{df}+1}{2}}}$$

where df is the number of degrees of freedom, in this case 14 and $\Gamma(n) = (n-1)!$.

```
x <- c(3.24,4.39,5.24,3.83,3.50,3.75,4.06,3.83,3.54,3.20,4.28,3.65,
      3.01,4.69,3.32)
mu <- 3.5
mux <- mean(x)
sigmax <- sd(x)/sqrt(length(x))
t <- (mux-mu)/sigmax
t
n=10000
arr <- seq(0, t, length=n)
fsum=0
i=0
df=14
for(x in arr){
fsum=fsum+(3-(-1)^i)*(gamma((df+1)/2)/(sqrt(pi*df)*gamma(df/2)*
  (1+x^2/df)^((df+1)/2)))
i=i+1
}
fsum*((arr[2]-arr[1])/3)
rs=fsum*((arr[2]-arr[1])/3)
r=rs*2
r1=1-r
p=r1/2
p
```

Outputs:

```
> t
[1] 2.129865
> p
[1] 0.02566876
```

- Assuming the same significance level of 0.05, we reject the null hypothesis since $p = 0.0257 < 0.05$.
- A t-test is stricter than a Z-test whenever the sample size is small. In our case, the sample size is 15, which is small and hence the results of a t-test are closer to the actual scenario than a Z-test. For higher sample sizes, t and z-test results are approximately the same.

Exercise 11.3

- The pooled-variance t-test assumes that the variances of the populations in the two groups under test are equal. The pooled variance attempts to estimate the overall population

variance from two or more samples taken from the same population. When sample sizes for the t-test are small, the variance estimates for the individual samples may not be accurate and a higher precision is obtained by estimating a common variance for the entire population by pooling in data from the two samples.