hw4

Thomas Le, ID: 91308173 and Armand Nasseri, ID: 912679383

June 9, 2019

1

a)

Supposing process begins at state 1 at T=1. Probability it will be at state 0 at t=3 is the sum of:

```
1 -> 0 -> 0 = .7 * .4 = 0.28
1 -> 1 -> 0 = .3 * .7 = 0.21

Then, the answer is 0.28 + 0.21 = 0.49
```

b)

Solving for steady-state distribution solving for $\mu P = \mu$ using eigenvalues

Source: https://nicolewhite.github.io/2014/06/10/steady-state-transition-matrix.html

```r
# markov chain table setup
p = matrix(c(0.4, 0.7,
             0.6, 0.3),
           nrow = 2,
           ncol = 2)

states = c(0, 1)
colnames(p) = states
rownames(p) = states
p
```

```
##     0   1
## 0 0.4 0.6
## 1 0.7 0.3
```

```r
# solving using eigen values
e = eigen(t(p))
first = Re(e$vectors[, 1])

# normalize
mu = first / sum(first)

mu
```

```
## [1] 0.5384615 0.4615385
```

2

Given mean a Poisson distribution with mean of 4.2 bugs per 1,000 lines of code. Searching for probability of more than 40 bugs in 20 sections of code, each 1,000 lines long, where sections are independent.

For a Poisson distribution, E[X] = λ. Given that the mean is 4.2, then $\lambda = 4.2$

For a set of 20 sections, then the mean is now `E[X] = 4.2 * 20 = 84`.

Similarily, `Var[X] = 4.2 * 20 = 84`.

Then, for a set of i.i.d random variables, its distribution is approximate normal, $N(84, \sqrt{84})$.

Then, P(X > 40) = P(z > $\frac{40-84}{\sqrt{84}}$) = P(z > -4.80). Though -4.8 is not on the z table, the z value is very far to the left of the distribution so the p value should be approximate 0.0001. Then the probability is $1 - 0.0001 = 0.9999$.

The r function to find the z value would be $pnorm(40, 84, \sqrt{84})$.

So the answer using R would be $1 - pnorm(40, 84, \sqrt{84}) = 1 - 7.933e\text{-}07 = 0.9999$

3

Given an exponential distribution with mean = 15, then because for an exponential distribution, $E[X] = \frac{1}{\lambda}$, then $\lambda = \frac{1}{15}$.

```r
# generate 10 random numbers
lambda <- 1/15
weights <- rexp(10, lambda)
cat("Randomized weights are:")
```

```
## Randomized weights are:
```

```r
weights
```

```
##  [1] 28.211624 19.313236  3.602066 13.960602 27.526326 65.461260 42.457753
##  [8]  1.041707 12.739653 15.696580
```

```r
z <- qnorm(0.93)

xbar <- mean(weights)

s2 <- mean(weights^2) - xbar^2
s <- sqrt(s2)
radius <- z * s/sqrt(10)
cat("Approximate confidence interval is", xbar-radius, "to", xbar+radius, "\n")
```

```
## Approximate confidence interval is 14.47007 to 31.53209
```

4

To test that the length of the two fish (salmon S, tuna T) have no difference against $\alpha = 0.05$, let:

$$H_0 : \overline{S} - \overline{T} = 0$$
$$H_A : \overline{S} - \overline{T} \neq 0$$

```r
# salmon s
sbar <- 14
sstd <- 2
ssize <- 20

# tuna t
tbar <- 15
tstd <- 3
tsize <- 15

# standard error
se <- sqrt( ((sstd^2)/ssize) + ((tstd^2)/tsize))

Z <- (sbar-tbar-0)/se

# p value
pval <- 2*pnorm(Z)
pval
```

```
## [1] 0.2635525
```

The P-value for the test statistic is 0.2635525, which is not less than the alpha value of 0.05.

Thus, we fail to reject the null hypothesis at the 0.05 significance level.

In context, we fail to reject the hypothesis that there is no difference in the length of salmon vs tuna in the sacramento river.