IS604_charleyferrari_hw7

Charley Ferrari

November 21, 2015

Question 9.14

The highway between Atlanta and Athens has a high incidence of accidents along its 100 kilometers. Public safety officers say that the occurrence of accidents along the highway is randomly (uniformly) distributed, but the news media say otherwise. The Georgia Department of Public Safety published records for the month of September. These records indicated the point at which 30 accidents involving an injury or death occurred, as follows (the data points representing the distance from the city limits of Atlanda)

H_0: The data is uniformly distributed between 0 and 100.

H_a: The dat ais not uniformly distributed between 0 and 100

```r
accidents <- c(88.3,40.7,36.3,27.3,36.8,91.7,67.3,7,45.2,23.3,98.8,90.1,17.2,
               23.7,97.4,32.4,87.8,69.8,62.6,99.7,20.6,73.1,21.6,6,45.3,76.6,
               73.2,27.3,87.6,87.2)

l <- 100

accidentsks <- accidents/l

dplus <- max((1:length(accidentsks))/length(accidentsks) - accidentsks)

dminus <- max(accidentsks -
                ((1:length(accidentsks))-1)/length(accidentsks))

d <- max(c(dplus,dminus))

d
```

```
## [1] 0.883
```

According to table A8, if we're looking at a p-value of 0.05, the critical value of D is 0.24. Our test statistic D is indeed larger than this, so we can reject the null hypothesis that these accidents are uniformly distributed.

Question 9.17

The time required for 50 different employees to compute and record the number of hours worked duringthe week was measured, with the following results in minutes.

Use the chi-square test to test the hypothesis that these service times are exponentially distributed. Let the number of class intervals be k = 6. Use the level of significance $\alpha = 0.05$.

```r
times <- c(1.88,0.54,1.9,0.15,0.02,2.81,1.5,0.53,2.62,2.67,3.53,0.53,1.8,0.79,
           0.21,0.8,0.26,0.63,0.36,2.03,1.42,1.28,0.82,2.16,0.05,0.04,1.49,0.66,
           2.03,1,0.39,0.34,0.01,0.1,1.1,0.24,0.26,0.45,0.17,4.29,0.8,5.5,4.91,
           0.35,0.36,0.9,1.03,1.73,0.38,0.48)
```

$a_i = -\frac{1}{\lambda} ln(1 - ip)$, for $i = 0, 1, ..., k$

$\hat{\lambda} = \frac{1}{\bar{X}}$

```
k <- 6

lambda <- 1/mean(times)

p <- 1/k

ai <- -(1/lambda)*log(1-(0:k)*p)

O <- length(times[times<ai[2]])

i <- 2

for(i in 2:6){
  O <- c(O,length(times[times >= ai[i] & times < ai[i+1]]))
}

E <- rep(50/k,k)

chisq <- sum(((O-E)^2)/E)
```

s = 1 for the exponential distribution, so our degrees of freedom are k-s-1 = 6-1-1 = 4

```
qchisq(0.95,4)
```

```
## [1] 9.487729
```

```
chisq
```

```
## [1] 2.8
```

Our test statistic is well outside the acceptable range, so we can reject the null hypothesis. This data doesn't appear to be exponentially distributed.

Question 10.1

A simulation model of a job shop was developed to investigate different scheduling rules. To validate the model, the scheduling rule currently used was incorporated into the model and the resulting output was compared against observed system behavior. By searching the previous year's database records, it was estimated that the average number of jobs in the shop was 22.5 on a given day. Seven independent replications of the model were run, each of 30 days duration, with the following results for average number of jobs in the shop:

```
times <- c(18.9, 22, 19.4, 22.1, 19.8, 21.9, 20.2)
```

a. Develop and conduct a statistical test to evaluate whether model output is consistent with system behavior. Use level of significance $\alpha = 0.05$

H_0: $\mu = 22.5$ jobs

H_a: $\mu \neq 22.5$ jobs

$$t_0 = \frac{\bar{Y} - \mu_0}{S/\sqrt{n}}$$

```
mu <- 22.5

t0 <- (mean(times)-mu)/(sd(times)/sqrt(length(times)))

t0
```

```
## [1] -3.679941
```

This is a two sided test, and our t-statistic is negative. So the critical value of t is found below:

```
qt(0.025,length(times)-1)
```

```
## [1] -2.446912
```

Our test statistic is outside the acceptable range, so we reject the null hypothesis that the mean number of jobs in our model is 22.5.

b. What is the power of this test if a difference of two jobs is viewed as critical? What sample size is needed to guarantee a power of 0.8 or higher (using $\alpha = 0.05$)?

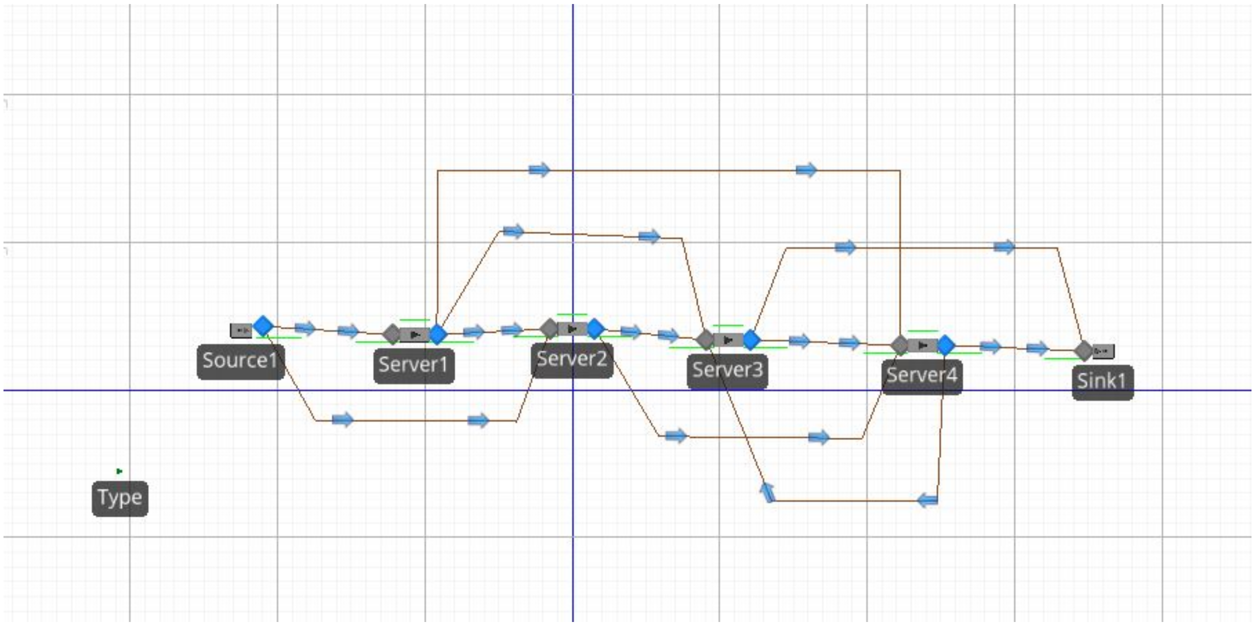
β, the probability of a type II error, depends on δ:

```
delta <- abs(mean(times)-mu)/sd(times)
```

This is a two-sided test, so $\beta(\hat{\delta})$ can be calculated from table A.10. $\beta_{n=7}(1.4) = 0.1$. Therefore, $1 - \beta = 0.9$.

We can't tell for sure from this table what the n would be, but it appears to be between 5 and 7. I'd guess 6 is the minimum sample size needed to get a power of 0.8. If the sample size were 5 in this case, assuming the same sample standard deviation (which is also dependent on n…), n=5 would give a β around 0.3, or a power of 0.7.

Question 11.13



This Simio model is attached to the homework. The "type" ModelEntity has its initial priority defined according to the given proportions using the random.discrete function. I created a data table, defining the service times depending on the ModelEntity type as follows:

	Entity Priority	Station1Time (Hours)	Station2Time (Hours)	Station3Time (Hours)	Station4Time (Hours)
1	1	Random.Normal(20,3)	Random.Normal(30,5)	Random.Normal(75,4)	Random.Normal(20,3)
2	2	Random.Normal(18,2)	0.0	Random.Normal(60,5)	Random.Normal(10,1)
3	3	0.0	Random.Normal(20,2)	Random.Normal(50,8)	Random.Normal(10,1)
4	4	Random.Normal(30,5)	0.0	0.0	Random.Normal(15,2)

And defining the path weights depending on the ModelEntity type as follows:

	n4Time (Hours)	Path12Prob	Path23Prob	Path34Prob	Paths1Prob	Path13Prob	Path4s Prob	Paths2Prob	Path24Prob	Path43Prob	Path3s Prob	Path14Prob
1	om.Normal(20,3)	1	1	1	1	0	1	0	0	0	0	0
2	om.Normal(10,1)	0	0	1	1	1	1	0	0	0	0	0
3	om.Normal(10,1)	0	0	0	0	0	0	1	1	1	1	0
4	om.Normal(15,2)	0	0	0	1	0	1	0	0	0	0	1

The path weights change from 1 or 0 depending on the type of the ModelEntity.

Here's a table of the average worker utilization and mean total response time for each job type: