Statistics 140 Winter 17

Hand-In Assignment #8

Sarah Ruckman

Last 4 Digits of SID: 7194

1. If you subscribe to a cell phone service, how many numbers do you have on your plan? According to the Current Population Survey (CPS) Cell Phone Supplement, 49% of cell phone plans have only 1 number, 39% have 2 numbers, 8% have 3 numbers and 4% have 4 or more numbers. Brandon recently obtained a random sample of 500 Riverside cell phone users and recorded the following data:



Perform the appropriate test of hypothesis using α = 0.05 to see if the Riverside area agrees with the national survey. (7 pts)

**P = number of users**

**H0: The data follows a multinomial distribution with certain proportions**

**(p1 = 0.49, p2 = 0.39, p3 = 0.08, p4 = 0.04)**

**Ha: The data does not follow a multinomial distribution with certain proportions**

**(p1 ≠ 0.49, p2 ≠ 0.39, p3 ≠ 0.08, p4 ≠ 0.04)**

SAS Code:

options ls = 78 ps = 55 formdlim = '#' nocenter nodate;

ods graphics off;

data q1;

input users $ observed1 prop1;

n = 500;

expected1 = n\*prop1;

chisq1 = (observed1 - expected1)\*\*2/expected1;

datalines;

a 245 0.49

b 200 0.39

c 45 0.08

d 10 0.04

;

proc print noobs;

proc means sum noprint;

var chisq1;

output out = a1 sum = chisq\_sum;

proc print noobs;

data pvalue1;

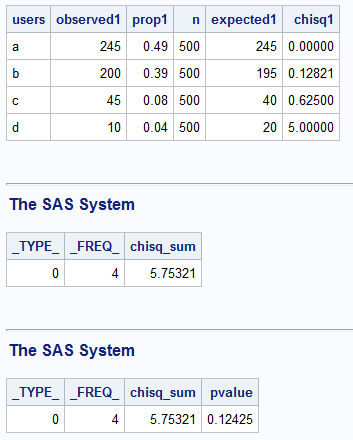
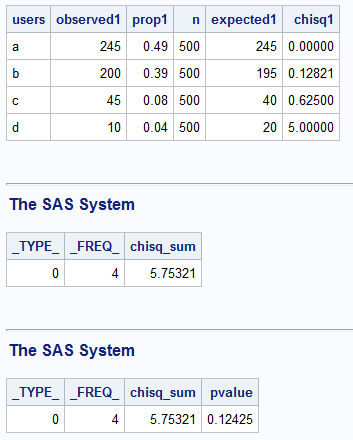
set a1;

pvalue = 1 - probchi(chisq\_sum,3);

proc print noobs;

run;

quit;



**TS: X2 = 5.75321 with p-value = 0.12425**

**RR: Reject H0 if p-value < α = 0.05**

**Since the p-value = 0.12425 is greater than α = 0.05, we do not reject H0**

**There is sufficient evidence to indicate that the ratio is correct. (p1 = 0.49, p2 = 0.39, p3 = 0.08, p4 = 0.04)**

1. A statistician, with Icelandic heritage1, was interested in examining the eruption height of Strokkur, a world-famous geyser in Iceland. 2 In particular, the statistician was interested in determining whether the eruption height was normally distributed. She obtained the following data: (The data has been saved in height1m w17.mtw and height1s w17.dat.)



1. Use Lilliefors Test to determine whether one can conclude that the geyser heights follow a normal distribution. Use α = 0.05. (6 pts)

**H0: The geyser heights are normally distributed**

**Ha: The geyser heights are not normally distributed**

R Code:

> g<-read.table("C:\\Users\\Sarah\\Downloads\\HEIGHT1S\_w17.dat", header=TRUE)

> g #Print as check

> attach(g)

> names(g)

[1] "height"

> mu\_est= mean(height)

> mu\_est

[1] 192.4

> sd\_est = sd(height)

> sd\_est

[1] 16.02221

> ks.test(height, "pnorm", mean= mu\_est, sd= sd\_est)

One-sample Kolmogorov-Smirnov test

data: height

D = 0.18237, p-value = 0.8363

alternative hypothesis: two-sided

**TS: D= 0.18237 with p-value = 0.8363**

**RR: Reject H0 if p-value < α = 0.05**

**Since the p-value = 0.8363 is greater than α = 0.05, we do not reject H0**

**There is sufficient evidence to indicate that the geyser heights follow a normal distribution.**

1. Use the Kolmogorov-Smirnov Test to determine whether one can conclude that the geyser heights follow a normal distribution with a mean of 200 feet and a standard deviation of 15 feet. (6 pts)

**H0: The geyser heights follow a normal distribution with a mean of 200 feet and standard deviation of 15 feet.**

**Ha: The geyser heights do not follow a normal distribution with a mean of 200 feet and standard deviation of 15 feet.**

R Code:

> ks.test(height,"pnorm", mean = 200, sd = 15)

One-sample Kolmogorov-Smirnov test

data: height

D = 0.30879, p-value = 0.2413

alternative hypothesis: two-sided

**TS: D = 0.30879 with p-value = 0.2413**

**RR: Reject H0 if p-value < α = 0.05**

**Since the p-value = 0.2413 is greater than α = 0.05, we do not reject H0**

**There is sufficient evidence to indicate that the geyser heights follow a normal distribution with mean of 200 feet and standard deviation of 15 ft.**

1. Create a graph of the empirical distribution function and the cumulative distribution function (Do not specify a mean or standard deviation.). (Be sure the two curves are overlaid on the same graph.) (2 pts)
2. Brandon, a quality control engineer claims that the life length of a particular light bulb for an LCD projector follows an exponential distribution with β = 1000. Rodrigo obtains a random sample of the 12 light bulbs and places them on life test with the following results (measured in hours). (The data has been saved in life1m.mtw and life1s.dat.)



1. Use Lilliefors Test to determine whether the life length of the bulb follows an exponential distribution. Use α = 0.05. (6 pts)

**H0: The life lengths are exponentially distributed.**

**Ha: The life lengths are not exponentially distributed.**

R Code:

> life1<-read.table("C:\\Users\\Sarah\\Downloads\\life1s.dat", header=TRUE)

> life1 #Print as check

> attach(life1)

> names(life1)

[1] "life"

> mu\_est = mean(life)

> mu\_est

[1] 1039.955

> ks.test(life1, "pexp", rate=1/mu\_est)

One-sample Kolmogorov-Smirnov test

data: life1

D = 0.18539, p-value = 0.7392

alternative hypothesis: two-sided

**TS: D = 0.18539 with p-value = 0.7392**

**RR: Reject H0 if p-value < α = 0.05**

**Since the p-value = 0.7392 is greater than α = 0.05, we do not reject H0**

**There is sufficient evidence to indicate that the life lengths are exponentially distributed.**

1. Use the Kolmogorov-Smirnov Test to determine whether the life length of the bulb follows an exponential distribution with a mean (i.e., β) of 1000 minutes. Use α = 0.05. (6 pts)

**H0: The life lengths are exponentially distributed with β = 1000.**

**Ha: The life lengths are not exponentially distributed with β = 1000.**

R Code:

> ks.test(life1,"pexp",rate=1/1000)

One-sample Kolmogorov-Smirnov test

data: life1

D = 0.1934, p-value = 0.6924

alternative hypothesis: two-sided

**TS: D = 0.1934 with p-value = 0.6924**

**RR: Reject H0 if p-value < α = 0.05**

**Since the p-value = 0.6924 is greater than α = 0.05, we do not reject H0**

**There is sufficient evidence to indicate that the life lengths are exponentially distributed with β = 1000.**

1. Create a graph of the empirical distribution function and the cumulative distribution function (do not specify a mean). (Be sure the two curves are overlaid on the same graph.) (2 pts)
2. Brandon, an outdoor enthusiast, was interested in examining the tensile strength of two brands of rock climbing ropes. He obtained nine rock climbing ropes from each of the manufacturers and measured the tensile strength of each rope. He recorded the following data: (The data has been saved in ropes1m.mtw and ropes1s.dat.)



1. Use the Kolmogorov-Smirnov Test to determine whether one can conclude there is a significant difference in tensile strength distribution functions between the two brands of rope. (In other words, test whether one can conclude there is a significant difference in the tensile strength between the two brands of rope.) Use α = 0.05. (6 pts)

**H0: Distribution functions of tensile strength are the same for the two brands of rope.**

**Ha: Distribution functions of tensile strength are not the same for the two brands of rope.**

R Code:

> a<-c(89.8,90.2,98.1,91.2,88.9,90.3,99.2,94,88.7)

> b<-c(87.3,76,66.7,77.3,86.4,86.4,93.1,89.2,90.1)

> ks.test(a,b)

Two-sample Kolmogorov-Smirnov test

data: a and b

D = 0.66667, p-value = 0.03663

alternative hypothesis: two-sided

Warning message:

In ks.test(a, b) : cannot compute exact p-value with ties

**TS: D = 0.66667 with p-value = 0.03663**

**RR: Reject H0 if p-value < α = 0.05**

**Since the p-value = 0.03663 is less than α = 0.05, we reject H0**

**There is sufficient evidence to indicate that the distribution functions of tensile strength are not the same for both brands of rope.**

1. Repeat part (i) using the Cramer-von Mises Test. (6 pts)

**H0: Distribution functions of tensile strength are the same for the two brands of rope.**

**Ha: Distribution functions of tensile strength are not the same for the two brands of rope.**

SAS Code:

options ls= 78 ps= 55 nocenter nodate;

ods graphics off;

data rope;

input brand strength @@;

datalines;

1 89.8 1 90.2 1 98.1 1 91.2 1 88.9 1 90.3 1 99.2 1 94 1 88.7

2 87.3 2 76 2 66.7 2 77.3 2 86.4 2 86.4 2 93.1 2 89.2 2 90.1

;

proc sort;

by brand strength;

proc print;

proc npar1way edf;

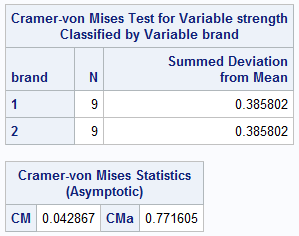
class brand;

var strength;

exact ks;

run;

quit;



**TS: CMa = 0.771605**

**RR: Reject H0 if CMa > 0.461**

**Since the CMa = 0.771605 is greater than 0.461, we reject H0**

**There is sufficient evidence to indicate that the distribution functions of tensile strength are not the same for both brands of rope.**

1. Create a graph of the two empirical distributions functions associated with the tensile strengths of the two brands of rope. (Be sure they are overlaid on the same graph!) (3 pts)