**Assignment #3**

**Stats 147 Fall 2017 Sec. 2**

Sarah Ruckman

SID: 7194

**Using R:**

1. A production process that creates specialized lighted dog leashes operates with 7% nonconforming (defective) output. A sample of 15 lighted dog leashes is selected at random and the number of nonconforming leashes counted.

Let Y = number of nonconforming lighted dog leashes. Then Y ∼ b(y; n = 15, p = 0.07). Use this information to complete the following

1. (1 pt) Generate the sequence 0 through 15 and store the sequence in the variable, y. (Be sure to print the values of y.)

**R Code:**

> #Stats 147 Assignment #3

> #Section 002

> #Fall 2017

> #Sarah Ruckman

> #R Question 1 part i

> #Generate a sequence using seq()

> #Call it y

> y <- seq(0,15)

> #Print as check

> y

[1] 0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15

1. (1 pt) Generate the probability distribution function of Y and store the values the variable pdf1. (Be sure to print the values of pdf1.)

**R Code:**

> #R Question 1 part ii

> #Generate pdf of Y and store it in pdf1 using dbinom function

> pdf1 <- dbinom(y,15,0.07)

> #Print as check

> pdf1

[1] 3.367009e-01 3.801461e-01 2.002920e-01 6.532823e-02 1.475154e-02

[6] 2.442727e-03 3.064353e-04 2.965503e-05 2.232099e-06 1.306725e-07

[11] 5.901338e-09 2.019031e-10 5.065668e-12 8.798927e-14 9.461212e-16

[16] 4.747562e-18

1. (1 pt) Generate the cumulative distribution function of Y and store the values in the variable cdf1. (Be sure to print the values of cdf1.)

**R Code:**

> #R Question 1 part iii

> #Generate cdf of Y and store it in cdf1 using pbinom function

> cdf1 <- pbinom(y,15,0.07)

> #Print as check

> cdf1

[1] 0.3367009 0.7168470 0.9171390 0.9824673 0.9972188 0.9996615 0.9999680

[8] 0.9999976 0.9999999 1.0000000 1.0000000 1.0000000 1.0000000 1.0000000

[15] 1.0000000 1.0000000

1. (2 pts) Generate the values of P(Y > y) and store the values in the variable sdf1. (Be sure to print the values of sdf1.)

**R Code:**

> #R Question 1 part iv

> #Generate the sdf or complement of cdf by subtracting 1

> #Call it sdf1

> sdf1 <- 1-cdf1

> #Print as check

> sdf1

[1] 6.632991e-01 2.831530e-01 8.286095e-02 1.753272e-02 2.781187e-03

[6] 3.384593e-04 3.202391e-05 2.368880e-06 1.367809e-07 6.108395e-09

[11] 2.070577e-10 5.154543e-12 8.892886e-14 8.881784e-16 0.000000e+00

[16] 0.000000e+00

1. (1 pt) Use the cbind function to generate a data set with y1, pdf1, cdf1 and sdf1 as the columns. (Be sure to print the data set.)

**R Code:**

> #R Question 1 part v

> #Use cbind to combine y, pdf1, cdf1, sdf1

> cbind(y,pdf1,cdf1,sdf1)

y pdf1 cdf1 sdf1

[1,] 0 3.367009e-01 0.3367009 6.632991e-01

[2,] 1 3.801461e-01 0.7168470 2.831530e-01

[3,] 2 2.002920e-01 0.9171390 8.286095e-02

[4,] 3 6.532823e-02 0.9824673 1.753272e-02

[5,] 4 1.475154e-02 0.9972188 2.781187e-03

[6,] 5 2.442727e-03 0.9996615 3.384593e-04

[7,] 6 3.064353e-04 0.9999680 3.202391e-05

[8,] 7 2.965503e-05 0.9999976 2.368880e-06

[9,] 8 2.232099e-06 0.9999999 1.367809e-07

[10,] 9 1.306725e-07 1.0000000 6.108395e-09

[11,] 10 5.901338e-09 1.0000000 2.070577e-10

[12,] 11 2.019031e-10 1.0000000 5.154543e-12

[13,] 12 5.065668e-12 1.0000000 8.892886e-14

[14,] 13 8.798927e-14 1.0000000 8.881784e-16

[15,] 14 9.461212e-16 1.0000000 0.000000e+00

[16,] 15 4.747562e-18 1.0000000 0.000000e+00

1. (1 pt) Find the probability that exactly 3 of the 15 lighted dog leashes are nonconforming (defective).

**R Code:**

> #R Question 1 part vi

> #Prob that exactly 3 are defective

> #Use dbinom to get prob of mass function

> p6 = dbinom(3,15,0.07)

> #Print as check

> p6

[1] 0.06532823

**The probability is 0.06532823.**

1. (1 pt) Find the probability that at most 3 of the 15 lighted dog leashes are nonconforming (defective).

**R Code:**

> #R Question 1 part vii

> #Find the prob of at most 3 or P(x<=3)

> p7 = pbinom(3,15,0.07)

> #Print answer

> p7

[1] 0.9824673

**The probability is 0.9824673.**

1. (2 pts) Find the probability that between, and including, 2 and 5 of the 12 lighted dog leashes are nonconforming (defective).

**R Code:**

> #R Question 1 part viii

> #Find the prob that is between 2 and 5 P(X<=5) - P(X<=1)

> p8 = pbinom(5,15,0.07) - pbinom(1,15,0.07)

> #Print the answer

> p8

[1] 0.2828145

**The probability is 0.2828145.**

1. (2 pts) Generate a plot of y vs pdf1. Be sure to give your plot an appropriate title.

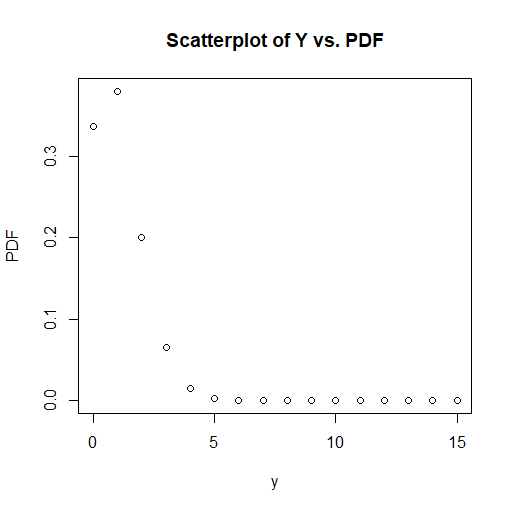
**R Code:**

> #R Question 1 part ix

> #graph y vs pdf1 give it a title!

> plot(y,pdf1,xlab="y",ylab="PDF",main="Scatterplot of Y vs. PDF")

> #the xlab and ylab function create the labels on those axises and main creates title



1. (2 pts) Generate a plot of y vs cdf1. Be sure to give your plot an appropriate title.

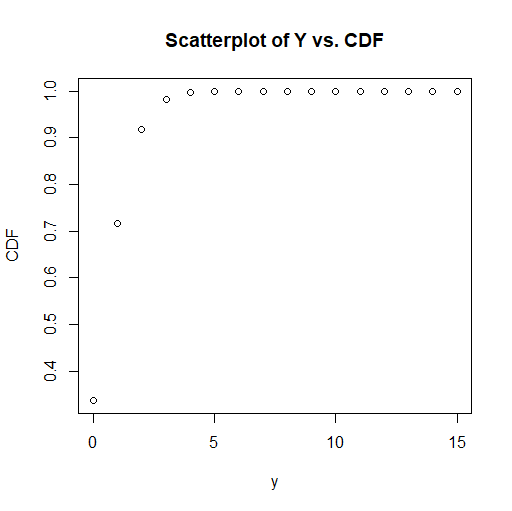
**R Code:**

> #R Question 1 part x

> #graph y vs cdf1 give it a title!

> #the xlab and ylab function create the labels on those axises and main creates title

> plot(y,cdf1,xlab="y",ylab="CDF",main="Scatterplot of Y vs. CDF")



1. Suppose the filling process of a particular type of dry dog food bag is normally distributed with a mean of 40 pounds and a standard deviation of 0.50 pounds. The filling process is considered to be functioning at an appropriate level (functioning “in control”) if the amount of fill in the dog food bags is between 38.75 pounds and 41.25 pounds. Lauren selects a dog food bag at random from the assembly line.
2. Find the probability that the bag is under-filled. (1 pt)

**R Code:**

> #R Question 2 part i

> #Find the prob that the bags are underfilled

> #mean = 40 sd = 0.5 underfilled = P(X<38.75)

> #Use pnorm function

> underfilled = pnorm(38.75,40,0.5)

> #Print answer

> underfilled

[1] 0.006209665

**The probability is 0.006209665.**

1. Find the probability that the bag is not under-filled. (1 pt)

**R Code:**

> #R Question 2 part ii

> #The prob that the bags are not underfilled

> #1-underfilled

> notunderfilled = 1-underfilled

> #Print answer

> notunderfilled

[1] 0.9937903

**The probability is 0.9937903.**

1. Find the probability that the bag is over-filled or under-filled. (1 pt)

**R Code:**

> #R Question 2 part iii

> #the prob that the bags are under or over filled

> #use pnorm lower = FALSE to find over filled

> over = pnorm(41.25,40,0.5,lower=FALSE)

> #Print answer

> over

[1] 0.006209665

> OR = underfilled + over

> #Print answer

> OR

[1] 0.01241933

**The probability is 0.01241933.**

1. Find the probability that the bag has a fill amount that meets specifications. (3 pts)

**R Code:**

> #R Question 2 part iv

> #Find the prob that it meets specs

> #P(38.75<=X=<41.25) = P(X<=41.25) - P(X<=38.75)

> between = pnorm(41.25,40,0.5) - pnorm(38.75,40,0.5)

> #Print answer

> between

[1] 0.9875807

**The probability is 0.9875807.**

1. Find the 97th percentile. (1 pt)

R Code:

> #R Question 2 part v

> #Use qnorm to find 97th percentile

> p97th = qnorm(0.97,40,0.5)

> #Print answer

> p97th

[1] 40.9404

**The 97th percentile is 40.9404.**

**Using SAS:**

1. Show Your School Spirit: Wednesday is R’Day. Staff, faculty, students and the community are being encouraged to show their UCR spirit by wearing blue and gold on Wednesday of each week. It has been estimated that 30% of UCR students are not aware of R’Day. A random sample of 26 UCR students was selected. Let X = # of students not aware of R’Day. Write a SAS program to complete the following.
2. (2 pts) Find the probability that exactly 10 of the selected students were not aware of R’Day?

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 1 Part i';

/\* Create temporary data set called students and input n=26, p=0.30, x1=10, x2=8, x3=13 \*/

**data** students;

input n p x1 x2 x3;

/\*use pdf to determine exactly 10 students\*/

p1 = pdf('Binom', x1,p,n);

datalines;

26 0.30 10 8 13

;

/\* Print the data only showing variables n p and p1 \*/

**proc** **print** noobs;

var n p p1;

**run**;

**quit**;

**The probability is 0.10424.**

1. (2 pts) Find the probability that no more than 10 of the selected students were not aware of R’Day?

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 1 Part ii';

/\* Create temporary data set called students and input n=26, p=0.30, x1=10, x2=8, x3=13 \*/

**data** students;

input n p x1 x2 x3;

/\*use pdf to determine exactly 10 students\*/

p1 = pdf('Binom', x1,p,n);

/\* use cdf to determine no more than 10 students \*/

p2 = cdf('Binom', x1,p,n);

datalines;

26 0.30 10 8 13

;

/\* Print the data only showing variables n p p1 and p2 \*/

**proc** **print** noobs;

var n p p1 p2;

**run**;

**quit**;

**The probability is 0.87471.**

1. (3 pts) Find the probability that between 9 and 13, inclusively, of the selected students were not aware of R’Day.

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 1 Part iii';

/\* Create temporary data set called students and input n=26, p=0.30, x1=10, x2=8, x3=13 \*/

**data** students;

input n p x1 x2 x3;

/\*use pdf to determine exactly 10 students\*/

p1 = pdf('Binom', x1,p,n);

/\* use cdf to determine no more than 10 students \*/

p2 = cdf('Binom', x1,p,n);

/\*use cdf of 13 - cdf of 8 to find value between \*/

p3 = cdf('Binom', x3,p,n)-cdf('Binom',x2,p,n);

datalines;

26 0.30 10 8 13

;

/\* Print the data only showing variables n p p1 p2 and p3 \*/

**proc** **print** noobs;

var n p p1 p2 p3;

**run**;

**quit**;

**The probability is 0.36316.**

1. (2 pts) Find the average number of students who were not aware of R’Day.

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 1';

/\* Create temporary data set called students and input n=26, p=0.30, x1=10, x2=8, x3=13 \*/

**data** students;

input n p x1 x2 x3;

/\*use pdf to determine exactly 10 students\*/

p1 = pdf('Binom', x1,p,n);

/\* use cdf to determine no more than 10 students \*/

p2 = cdf('Binom', x1,p,n);

/\*use cdf of 13 - cdf of 8 to find value between \*/

p3 = cdf('Binom', x3,p,n)-cdf('Binom',x2,p,n);

/\* use formula mean = n\*p to determine average \*/

mean1 = n\*p;

datalines;

26 0.30 10 8 13

;

/\* Print the data only showing variables n p p1 p2 and p3 \*/

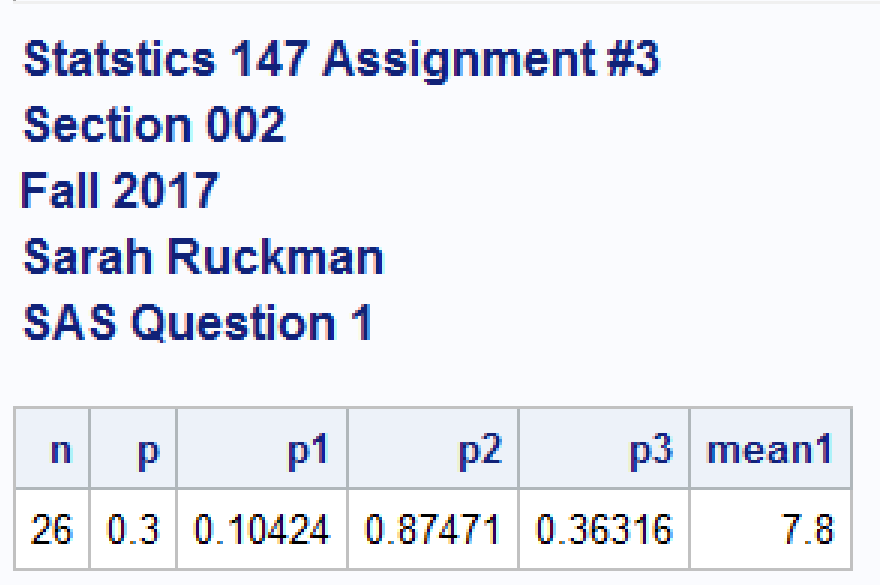
**proc** **print** noobs;

var n p p1 p2 p3;

**run**;

**quit**;

**The mean is 7.8**



1. A candy manufacturer claims that the weight of the candy in its bags of Halloween candy is normally distributed with a mean of 52 ounces and a standard of 3.6 ounces. The weight of the candy in the bag is considered an appropriate level (“in control”) if the weight of the candy in the bag is between 47.5 ounces and 55.5 ounces. Lauren selects a bag of candy from the assembly line. Let X = weight of the candy in the bag.
2. Find the probability that the weight of the candy in the bag is less than 50 ounces.

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 2 Part i';

/\* Create temporary data set called candy \*/

**data** candy;

/\* input variables mu=52 sigma=3.6 x1=50 x2=51 x3=55.5 x4=47.5 \*/

input mu sigma x1 x2 x3 x4;

/\* use cdf function to find prob that is less than 50\*/

p1 = cdf('Norm', x1,mu,sigma);

/\*Put in datalines\*/

datalines;

52 3.6 50 51 55.5 47.5

;

/\*Print the data and show only variables mu, sigma, and p1 \*/

**proc** **print** noobs;

var mu sigma p1;

**run**;

**quit**;

**The probability is 0.28926.**

1. Find the probability that the weight of the candy in the bag is more than 51 ounces

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 2 Part ii';

/\* Create temporary data set called candy \*/

**data** candy;

/\* input variables mu=52 sigma=3.6 x1=50 x2=51 x3=55.5 x4=47.5 \*/

input mu sigma x1 x2 x3 x4;

/\* use cdf function to find prob that is less than 50\*/

p1 = cdf('Norm', x1,mu,sigma);

/\* use sdf to determine more than 51\*/

p2 = sdf('Norm',x2,mu,sigma);

/\*Put in datalines\*/

datalines;

52 3.6 50 51 55.5 47.5

;

/\*Print the data and show only variables mu, sigma, p1, and p2 \*/

**proc** **print** noobs;

var mu sigma p1 p2;

**run**;

**quit**;

**The probability is 0.60941.**

1. Find the probability that the weight of the candy in the bag is at an appropriate level. (3 pt)

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 2 Part iii';

/\* Create temporary data set called candy \*/

**data** candy;

/\* input variables mu=52 sigma=3.6 x1=50 x2=51 x3=55.5 x4=47.5 \*/

input mu sigma x1 x2 x3 x4;

/\* use cdf function to find prob that is less than 50\*/

p1 = cdf('Norm', x1,mu,sigma);

/\* use sdf to determine more than 51\*/

p2 = sdf('Norm',x2,mu,sigma);

/\* Use cdf to find between 55.5 and 47.5 \*/

p3 = cdf('Norm',x3,mu,sigma) - cdf('Norm',x4,mu,sigma);

/\*Put in datalines\*/

datalines;

52 3.6 50 51 55.5 47.5

;

/\*Print the data and show only variables mu, sigma, p1, p2, and p3 \*/

**proc** **print** noobs;

var mu sigma p1 p2 p3;

**run**;

**quit**;

**The probability is 0.72888.**

1. Find the probability that the weight of the candy in the bag is not at an appropriate level. (1 pts)

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 2 Part iv';

/\* Create temporary data set called candy \*/

**data** candy;

/\* input variables mu=52 sigma=3.6 x1=50 x2=51 x3=55.5 x4=47.5 \*/

input mu sigma x1 x2 x3 x4;

/\* use cdf function to find prob that is less than 50\*/

p1 = cdf('Norm', x1,mu,sigma);

/\* use sdf to determine more than 51\*/

p2 = sdf('Norm',x2,mu,sigma);

/\* Use cdf to find between 55.5 and 47.5 \*/

p3 = cdf('Norm',x3,mu,sigma) - cdf('Norm',x4,mu,sigma);

/\* Use 1-p3 to determine out of control prob\*/

p4 = **1**-p3;

/\*Put in datalines\*/

datalines;

52 3.6 50 51 55.5 47.5

;

/\*Print the data and show only variables mu, sigma, p1, p2, p3, and p4 \*/

**proc** **print** noobs;

var mu sigma p1 p2 p3 p4;

**run**;

**quit**;

**The probability is 0.27112.**

1. Find the 98th percentile. (1 pt)

**SAS Code:**

options ls = **70** ps = **55** nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 2';

/\* Create temporary data set called candy \*/

**data** candy;

/\* input variables mu=52 sigma=3.6 x1=50 x2=51 x3=55.5 x4=47.5 \*/

input mu sigma x1 x2 x3 x4;

/\* use cdf function to find prob that is less than 50\*/

p1 = cdf('Norm', x1,mu,sigma);

/\* use sdf to determine more than 51\*/

p2 = sdf('Norm',x2,mu,sigma);

/\* Use cdf to find between 55.5 and 47.5 \*/

p3 = cdf('Norm',x3,mu,sigma) - cdf('Norm',x4,mu,sigma);

/\* Use 1-p3 to determine out of control prob\*/

p4 = **1**-p3;

/\* Use quantile function to find 98 percentile \*/

q98 = quantile('Norm', **0.98**,mu,sigma);

/\*Put in datalines\*/

datalines;

52 3.6 50 51 55.5 47.5

;

/\*Print the data and show only variables mu, sigma, p1, p2, p3, p4, and q98\*/

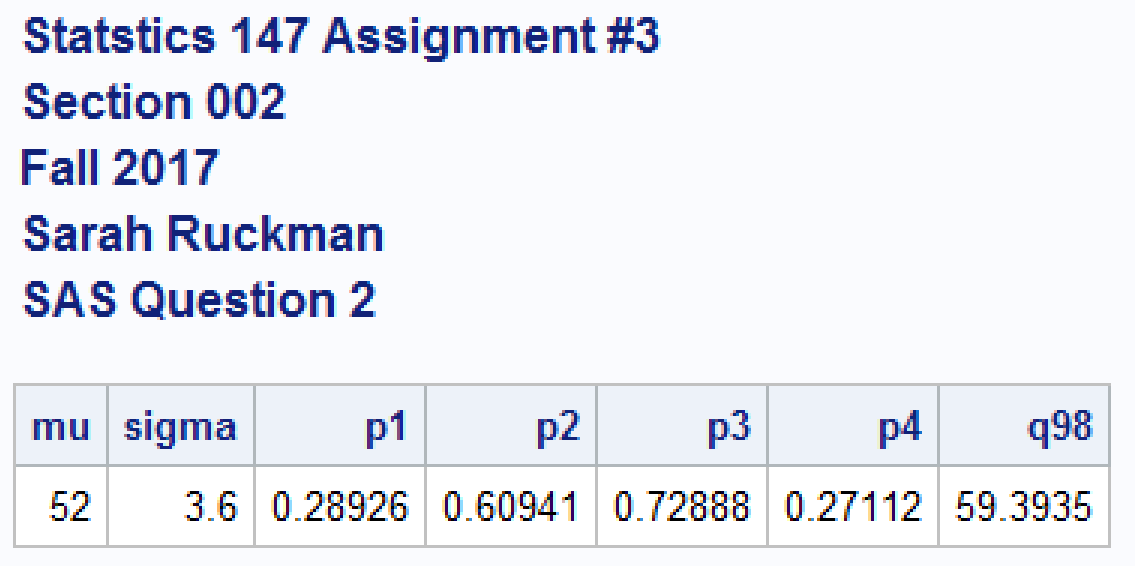
**proc** **print** noobs;

var mu sigma p1 p2 p3 p4 q98;

**run**;

**quit**;

**The 98th percentile is 59.3935.**



1. Let Let y = 3n 2 + √ 4.5m−0.5n for n = 1,3,5 and m = 2,5,8. Use nested DO loops to compute the values of y. (5 pts)

**SAS Code:**

options ls = 70 ps = 55 nocenter formdlim = '\*';

/\* ls = linesize, ps = pagesize, nocenter = justifies output, formdlim = overrides the internal page breaks

and replaces them with the designated symbol \*/

/\* Create titles \*/

title1 'Statstics 147 Assignment #3';

title2 'Section 002';

title3 'Fall 2017';

title4 'Sarah Ruckman';

title5 'SAS Question 3';

/\*Create temporary data set called loops\*/

data loops;

/\*Set up loop for n = 1,3,5 and m=2 to 8 in increments of 3 \*/

do n = 1 to 5 by 2;

do m = 2 to 8 by 3;

/\*Calculate y\*/

y = 3\*n\*\*2 + sqrt(4.5\*m\*\*(-0.5\*n));

/\*Output the information\*/

output;

/\*end the loops\*/

end;

end;

/\*Print the results\*/

proc print noobs;

run;

quit;

**Output/Answer:**

